Door detection in 3D coloured point clouds of indoor environments

Door Detection in 3D Coloured Point Clouds of Indoor Environments

B. Quintana, S. A. Prieto, A. Adán,
3D Visual Computing and Robotics Lab
Universidad de Castilla-La Mancha.
Ciudad Real, Spain
{Blanca.Quintana, Samuel.Prieto, Antonio.Adan}@uclm.es

F. Bosché
Centre of Excellence in Sustainable Building Design, CyberBuild Lab
Heriot-Watt University
Edinburgh, U.K.
f.n.bosche@hw.ac.uk

Abstract. *Door detection is becoming an increasingly important subject in building indoor modelling owing to its value in scan-to-BIM processes. This paper presents an original approach that detects open, semi-open and closed doors in 3D laser scanned data of indoor environments. The proposed technique is unique in that it integrates the information regarding both the geometry (i.e. XYZ coordinates) and colour (i.e. RGB or HSV) provided by a calibrated set of 3D laser scanner and a colour camera. In other words, our technique is developed in a 6D-space framework. The geometry-colour integration and other characteristics of our method make it robust to occlusion and variations in colours resulting from varying lighting conditions at each scanning location (e.g. specular highlights) and from different scanning locations. In addition to this paper, the authors also contribute a public dataset of real scenes along with an annotated ground truth. The dataset has varying levels of challenges and will help to assess the performance of new and existing contributions in the field. The approach proposed in this paper is tested against that dataset, yielding encouraging results.*

Keywords. ***Indoor spatial data model, 3D, point cloud, door detection, building information model, scan-to-BIM, robot, indoor navigation.***

Highlights.

- *New method for door detection in coloured 3D point clouds (6D data framework)*
- *The 6D data is obtained using a calibrated set of a laser scanner and an SLR camera with a flash*
- *The method is robust under conditions of occlusion and non-homogeneous illumination*
- *The method detects open, semi-open and closed doors.*
- *Performance is demonstrated with a dataset containing various levels of challenges made public by the authors.*

1 INTRODUCTION

Door detection is a critical functionality for automatic building scanning systems. For instance, autonomous mobile robots with 3D scanners must obtain precise information on the location and state of doors (open or closed) for robust and safe navigation (e.g. passing through doors) and manipulation (e.g. opening doors by grasping handles) [1,2]. Another application is the automated generation of as-is/as-built Building Information Models (BIMs) from laser scanned data – a process commonly called *Scan-to-BIM* – that requires the segmentation, recognition and precise positioning of all building components, including doors [3]. Door detection has become a necessary task in both of the contexts described above, and can be made even more difficult when clutter and occlusion conditions exist.

While the subject of door detection has been considered in previous research, this paper proposes a unique approach that:

(1) integrates both geometric and colour information, provided by a calibrated set of 3D laser scanner and a colour camera;

(2) ensures reliable colour information by (a) employing a camera flash to reduce colour variations resulting from non-homogeneous illumination conditions experienced at different scanning locations; (b) detecting and correcting specular highlights that often result from the use of the camera flash; and (c) optimally merging colour information by assessing the suitability of each scanning location as regards acquiring the colour of any part of the scene;

(3) presents a general solution for open, semi-open and closed doors, providing the opening angle;

(4) provides the accurate size and pose of each door in the 3D world-coordinate-system; and

(5) is robust to clutter in the room and the resulting occlusions of the walls.

As will be shown in the review of Related Works in Section 2, existing door detection methods have typically considered only one or two of those aspects.

The document is organised as follows. Section 2 provides a review of the state of the art in door detection in 3D environments. Section 3 sets the general context in which the paper has to be considered, in order to enable the reader to fully understand the inputs of our approach. Our proposed approaches for specular highlight detection and correction, in addition to multiple view merging, are described in Section 4. The door detection algorithm is presented in Sections 5 and 6. The experimental work and results are reported in the long Section 7. Section 8 deals with the choice of parameters and Section 9 presents the conclusion and proposes future improvements to the method.

2 RELATED WORK

Door detection in reality capture data (i.e. principally 2D or 3D imaging data) has already been studied for many years. This existing pool of prior research can be divided into two main approaches based on the type of data acquisition method considered: 2D colour imaging (using digital cameras) [4–9] and 3D imaging (using laser scanners or photogrammetric systems) [2,10–20].

Table 1 summarises the literature review by categorising the methods reviewed according to the input data considered and the applicability of the method (closed, open or semi-open doors). As can be seen, no approach has been proposed to date that has been shown to work with closed, open and semi-open doors. Our method, which is appended to the table, aims to achieve this by integrating both colour and 3D data. A detailed state of the art is presented in the following sub-sections.

Table 1. Categorisation of prior work on door detection in 2D colour and/or 3D data.

Method	Input data		Applicability		
	2D Colour	3D	Closed	Open	Semi-open
[6] [7] [8]	X		X		
[4] [5] [9]	X		X	X	
[10] [11]		X	X		
[1]		X		X	
[14][18][19]					
[13]		X		X	X
[2] [15] [16]	X	X	X		
[12][20]	X	X		X	
Ours	X	X	X	X	X

2.1 2D imaging-based methods

2D colour image-based approaches take advantage of the affordability of digital cameras. Furthermore, focusing on colour (instead of 3D) may be justified by the observation that doors are often within the wall plane, where 3D data may not provide significant added value. Yang and Tian [4] and Marwa M. Shalaby et al. [5] propose an approach based on the extraction of lines and corners in the colour image, and the subsequent detection of coherent sets of two horizontal and two vertical segments making up the door frame. Because they rely on the region boundary features (i.e. colour edges) instead of region features (which are more sensitive to changes in lighting and perspective), they are able to detect doors in challenging contexts, such as glass doors. However, when some of the door edges or corners are occluded (e.g. by curtains) both approaches might fail, and be unable to obtain the correct geometric door model. Yang and Tian's method [4] detects doors through the use of region boundary features, but does not distinguish between open and closed doors. This entails a lot of false positives (23%) in some of the challenging scenarios tested.

Andreopoulos et al. [6] present a method based on the two aforementioned approaches, which requires the door to be almost contained within the view of the camera. In addition to detecting the doorframe using geometrical features (i.e. corners and edges) in the image, this approach detects door handles using a learning algorithm trained with a large handle dataset (1,500 samples).

In contrast with the earlier methods, Chen et al. [7] detect doors by using a deep learning algorithm based on a convolutional neural network trained with a large number of examples. The problem with the false positives in approaches [4], [5] and [6] is solved here, but a large number of undetected doors (false negatives) appear. The performance of their method is shown only with closed doors, and no results are shown for open or half-open doors. With a somewhat reversed strategy to that of Yang et al. [4], Kim et al. [8] propose an approach that first detects door handles, and then uses them to claim the existence of doors in the scene. The method detects individual vertical lines and, assuming a specific height for the door handle,

obtains a RoI and determines the handle type. The authors prevent their approach from being usable to detect open doors.

Finally, Sekkal et al. [9] first generate a rough 3D model of the scene from the detection of vanishing lines in single colour images. This enables them to infer the location of wall planes within which they detect doors by looking for two consecutive vertical lines spaced by a predefined distance. This makes the method simpler but inefficient for the sizes of other doors. While the authors present some results that show that their ad-hoc method is able to detect both open and closed doors (without distinguishing them), it is certainly restricted to a small number of scenarios in which the images acquired actually contain the necessary vanishing lines (i.e. intersections of the walls with the floors and ceiling). This suggests that the method is not easily usable outside the context of corridor environments imaged with a front-facing camera.

The 2D image-based methods reviewed above are prone to produce large numbers of false positives owing to their typical lack of consideration for the structure of the scene (i.e. where walls are) and the presence of many objects that are rectangular like doors (e.g. windows, paintings, radiators, or furniture).

2.2 3D imaging-based methods.

In order to achieve higher precision and reliability, researchers are increasingly considering 3D reality capture sensors, sometimes together with 2D colour images [1,2,10–16,21]. This strategy is motivated by the value of 3D data in the understanding of the structure of the (indoor) environment, but also by the rapidly decreasing price of these sensors.

Goron et al. [10] obtain point clouds from a 2D Laser Range Finder (LRF) on a tilting platform and extract the planes corresponding to closed doors by applying RANSAC (Random Sample Consensus). The main limitation of this technique is that it is based on the strong assumption that door panels do not lie exactly on the wall planes, which is often not true. Using a 2D range camera, Meyer zu Borgsen et al. [11] segment the 3D point cloud of the scene into planar patches using a region-growing algorithm based on point normal vectors. A door is detected if and only if the dimensions of the detected plane match pre-defined ‘standard’ dimensions, and the door plane contains a handle. This approach is optimized to detect single-leaf closed doors and also assumes that door panels do not lie exactly on the wall planes.

Using only a depth sensor (i.e. no colour information), Derry and Argall [14] and Dai et al. [1] detect walls as vertical planes in the acquired point clouds and subsequently detect doors as gaps in the point clouds of wall planes. Xu et al. [18] and Budroni et al. [19] use the same principle referred above and detect open doors. Yuan et al. [13] extend this approach to the detection of open or half-open doors. Wall planes are extracted from the point clouds acquired by a depth camera and the door’s opening angle is calculated by analysing the shape of the gap inside the door. Although interesting, this approach requires that the sensor be placed exactly in front of the door.

In the following paragraphs, we provide a more in-depth discussion regarding four of the methods most closely related to ours. In all of them, as in ours, the authors use colour and 3D geometry to detect doors in robotic environments. We have selected those of Varadarajan et al. [12], Kakillioglu et al [20], Adiwahono et al. [2], Díaz-Vilariño et al. [15] and Banerjee et al. [16].

Varadarajan et al. [12] propose an approach focused on the 3D room modelling that detects open doors in 3D point clouds obtained with a stereo camera rig. Colour information is used to identify wall-like surfaces, being the vertical planes calculated by means of iteratively reweighted least squares robust linear regression. Doors are searched for as gaps in the point clouds of wall planes (i.e. regions on the wall without sensed data).

Kakillioglu et al. [20] also use 3D data and colour information in their approach. First, the planes corresponding to the walls are segmented by means of the RANSAC algorithm and the regions that contain gaps are then identified. Since the gaps may come from windows, mirrors or shiny surfaces, a verification stage is applied on a colour image which covers the gap's surrounding. In order to identify specifically gaps inside doors, they use a learning technique called Aggregate Channel Features (ACF). The method only detects open doors and does not provide the doorframe's 3D coordinates.

Adiwahono et al. [2] use a horizontal line scanner and, similarly to Goron et al. [10], assume that the door is not entirely flush with the wall. The door candidate is detected as a cluster of points forming a relatively straight line of a specified length. The robot then approaches each candidate door and scans it with a range camera. The handle is detected by matching the mesh model of the handle with the data. This method does not delimitate the boundary of the door and only works for closed doors.

Díaz-Vilariño et al. [15] carry out the detection of closed doors by applying the Generalized Hough Transform on RGB orthoimages of the wall extracted from coloured point clouds acquired with a laser scanning system. This method focuses on the detection of rectangles in the colour data, and is thus only able to detect closed doors. The approach is limited to cases in which the wall and door are different colours.

In the framework of the Darpa Robotic Challenge, Banerjee et al. [16] develop an approach that enables an Atlas robot to detect closed doors. Doors are detected by finding consecutive pairs of vertical lines at a specific distance from one another in a 2D colour image of the scene. The lines are then recalculated in a 3D space with the help of the RANSAC algorithm. If there is a flat surface between each pair of lines, it is recognised as a closed door. Handle detection is subsequently carried out by means of colour segmentation, on the assumption that the handle is a different colour from that of the door. This approach makes several important assumptions. Among others, the authors impose a specific size on the door and handle, demand different colours for the wall and door and require the door to be in front of the 3D sensor.

Table 2 shows other essential aspects that differentiate our approach from the works referenced above. The respective columns refer to the following aspects. (1) Wall detection in the door detection process; (2) Extraction of the door contour: some approaches only recognise the door handle ([2]) or the door opening ([12]) but do not delimitate the door’s boundaries, (3) Door type (C=closed, O=open, S=semi-open), (4) Integration of several views of the door, (5) Multiple door detection on the wall; (6) Door’s opening angle, (7) Dealing with varying light conditions (specular highlight) and colour variation; (8) Dealing with occlusion (9) Restrictions regarding the wall vs. door colours, (10) Door size restrictions.

Table 2. Comparison with the most related 3D imaging-based methods

Method	(1) Wall detection	(2) Door's contour	(3) Types	(4) Views integration	(5) Multiple doors	(6) Op. angle	(7) Dealing with colour variation	(8) Occlusion	(9) Restriction: Wall/Door colours	(10) Restriction: Door's size
[12]	X	-	O	-	X	-	X	-	-	X
[20]	X	-	O	-	X	-	-	-	-	X
[2]	-	-	C	-	-	-	-	-	-	X
[15]	X	X	C	X	X	-	-	-	X	X
[16]	-	X	C	-	-	-	-	-	X	X
Ours	X	X	C/O/S	X	X	X	X	X	-	X

This paper is an extended and improved version of an earlier publication [22], which was an initial and incomplete solution that was developed under some restrictions that no longer exist in the improved system reported here. The essential differences between the initial and current versions are as follows:

1. Our initial pipeline only worked under non-specular conditions. However, in order to reduce the impact of varying light conditions, our system makes use of a camera flash that has the side-effect of potentially generating large specular highlights that may have an impact on colour-based data processing strategies. The system reported here addresses this issue with an additional pre-processing stage, implementing a new specular highlight detector and corrector.
2. In the earlier work, colour and depth images from several views of the scene were manually integrated. In the new system, an automatic *view merging* solution is proposed that optimally integrates several views into a unique 4D RGB-D orthoimage for each wall, taking into account specular highlights and the value of each scanning location to define the colour of each part of the wall.
3. In this new work, we deal with doors in any state (open, semi-open and closed) and in fact present an integrated approach that provides the opening angle of the door.
4. More extended experiments are reported in this paper that are conducted using a new dataset that is larger in size (number of test cases), variety and complexity. Additionally, this new dataset, which includes detailed ground truth information, is made publically available [23].

3 OVERALL DATA ORGANISATION

The work presented here focuses on door detection, that is performed once the scanning of a room has been completed. The output of the room scanning is composed of (1) a dense 3D coloured point cloud; (2) a labelled voxel model with associated 3D points from the point cloud; and (3) a 3D boundary model of the room composed of planar rectangular patches (and their associated voxels) representing the walls, ceilings and floors. Figure 1 illustrates the room scanning process and the generation of the labelled voxel model. We have carried out experiments to discover that a 20cm/side voxel dimension provides a suitable trade-off between data size (and therefore processing time) and performance. For more detailed information on our system, and particularly the creation of the voxel model and voxel labelling, we direct the reader to our prior publication [24].

Since doors are located in walls, we shall from here on focus on these SEs. As shown in Figure 2, wall elements have associated voxels that can be labelled as either:

- *Occupied*: The voxel contains at least one scanned point.
- *Occluded*: The voxel does not contain any point and was not visible from any of the scanning locations used to scan the room.
- *Opening*: The voxel does not contain any point, despite being visible from at least one scanning location.

Our door detection process considers the labelling and coloured 3D points associated with the voxels of each wall rectangular segment.

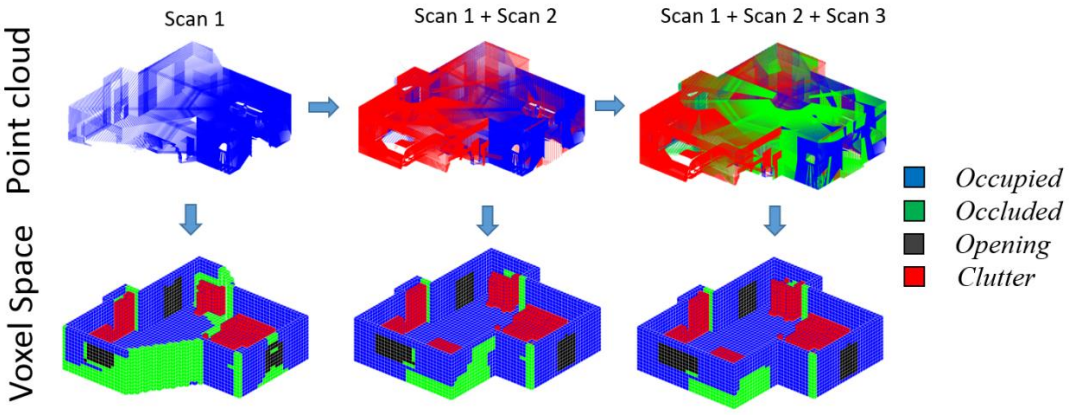


Figure 1. Illustration of the process used to construct the 3D voxel space and labels. Occupied voxels that are not associated with SEs are labelled as Clutter. Voxels belonging to the ceiling have been omitted for a better visualisation.

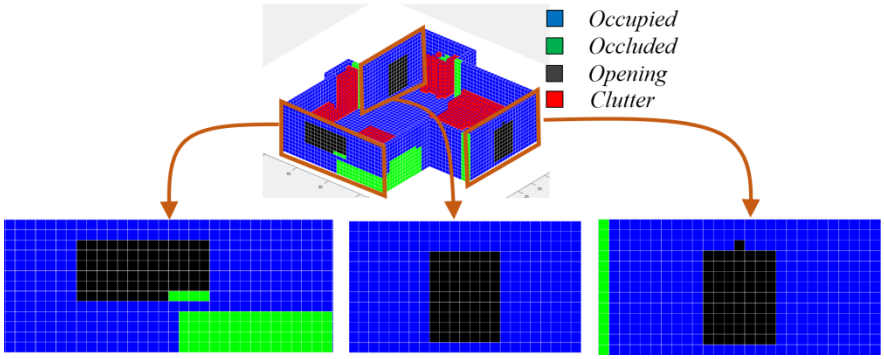


Figure 2. Front view of the labelled voxels of walls. Wall voxels can be labelled as Occupied, Occluded or Opening.

4 WALL DATA PREPARATION – VIEW MERGING

The proposed algorithm for door detection uses as input a *4D orthoimage* J_{CD} of the wall in which each pixel has colour (RGB or HSV) and depth (i.e. orthonormal distance of the 3D points to the wall plane). The resolution of the 4D RGB-D *orthoimage* is set much higher than that of the voxel space, with a pixel size of 5mm × 5mm.

The coloured 3D point clouds associated with each of the walls extracted are acquired from various scanning locations. Integrating this information into a single reliable 4D orthoimage requires a robust *view merging* approach that considers both *geometric merging* and *colour merging*.

The proposed view merging approach consists of creating, for each wall plane, a 4D orthoimage J_{CD}^{k} for each of the k scanning locations that contribute data (i.e. coloured 3D points) to that wall, and then merging those multiple 4D orthoimages into a unified orthoimage J_{CD}. Each 4D orthoimage J_{CD}^{k} has the same size as J_{CD}.

Owing to the discretization effect, several coloured 3D points could be contained in the same pixel of J_{CD}^{k} . For each orthoimage J_{CD}^{k}, multiple colours and depths may, therefore, be associated with any given pixel, which requires a first level of geometric and colour merging. For colour merging, we set the colour of the pixel to the mean of the colours of the 3D points associated with it. For the geometric merging, we set the depth of the pixel to the maximum of the depths of the 3D points, i.e. the depth of the point that is the furthest from the wall plane inside the room.

The merging of the set of orthoimages $\{J_{CD}^{k}\}$ into the unified J_{CD} orthoimage is then carried out for the geometric information independently from the colour information. For the geometric merging, we similarly set the pixel depth to the maximum of the depths of the same pixel in all orthoimages in $\{J_{CD}^{k}\}$. For the colour merging, we propose an original approach, described in more detail in the following sections, that takes into account the presence of specular highlights and the 'scanning value' of each scanning location to define the colour of each part of the wall.

Section 4.1 first presents our proposed approach for specular highlight detection. Section 4.2 then presents the proposed colour merging approach that considers the value of each scanning position in terms of both colour sensing and the specular highlights detected.

4.1 Detection of specular highlight regions

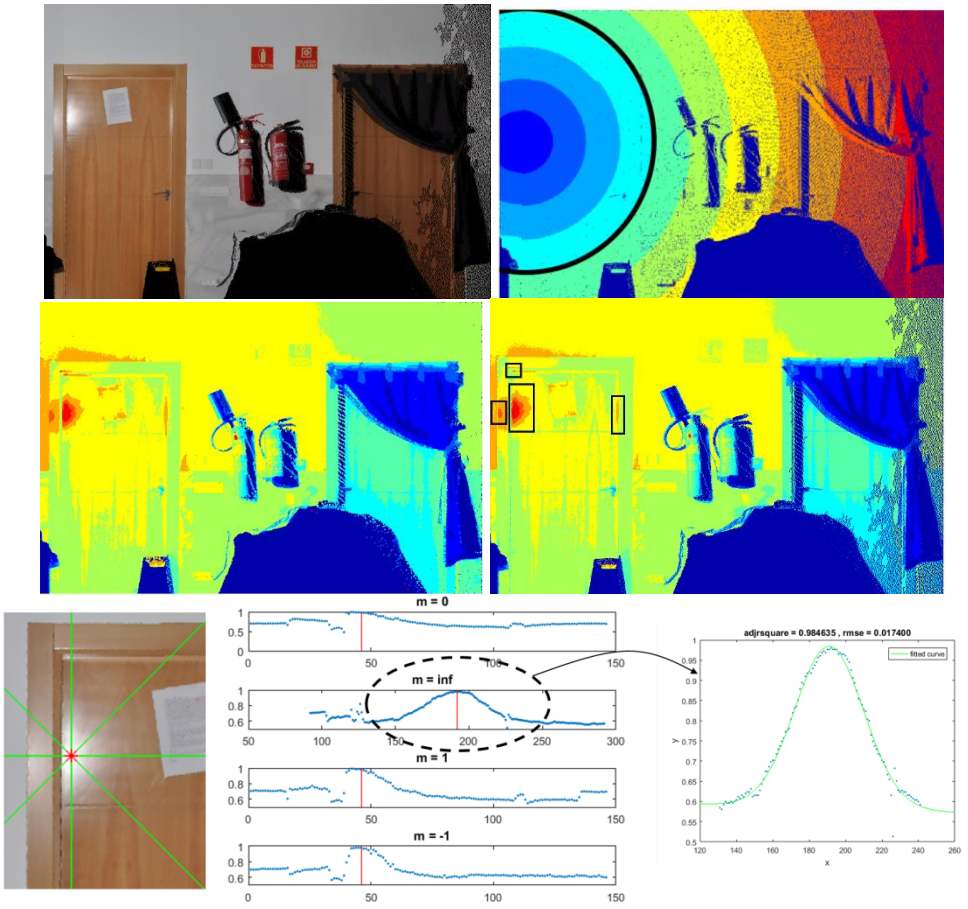
Since we deal with indoor environments with variable illumination conditions (due to natural light coming from windows and uncontrolled artificial lighting), the photos from the scanner's camera must be taken with flash. This provides a better, generally more consistent illumination of the scene; but at the same time this may result in specular highlights of various magnitude. To ensure that reliable colour information is provided for subsequent data processing stages (e.g. door detection), these specular highlights must be robustly detected and corrected. For specular highlight detection, the following four-step algorithm is proposed that is applied to each 4D orthoimage J_{CD}^{k} (see also illustration in Figure 3):

1. *Specular Highlight Region of Interest (RoI)* (Figure 3 (a)). Since the geometric data associated to walls is mainly planar, we make the simple assumption that specular highlights should mainly occur in regions surrounding the orthonormal projection of the camera centre on the wall surface. However, we do not assume that their extent is isotropic (i.e. they have a circular shape). Instead, we consider that their extent may be affected by the surface material and local surface geometry.
2. *Specular Highlight Region Candidates* (Figure 3 (b)). Candidate regions are detected in the wall's quantized intensity image with 11 grey levels. This quantized grey-scale image

is segmented and each segment analysed. If a segment is entirely surrounded by other lower-level segments and it does not contain any other segment, then we consider that it is a specular region candidate.

3. *Specular Highlight Region Detection* (Figure 3 (c)). For each candidate region, four intensity profiles along the North/South (N/S), East/West (E/W), NE/SW and NW/SE axes and passing through the region's centroid are analysed. If any of them fits a 2D Gaussian function, the entire region is recognized as a specular highlight region.

4. *Extent of Detected Specular Region* (Figure 3 (d)). The extent of each detected specular region is defined by eight points coming from the above profiles, each profile of the specular region providing two end-points. The end-points are found either where the slope of the profile from the maximum point reaches zero, or where the profile presents a strong discontinuity. Finally, the specular highlight region is delimited by the spline that takes the eight end-points as control points.



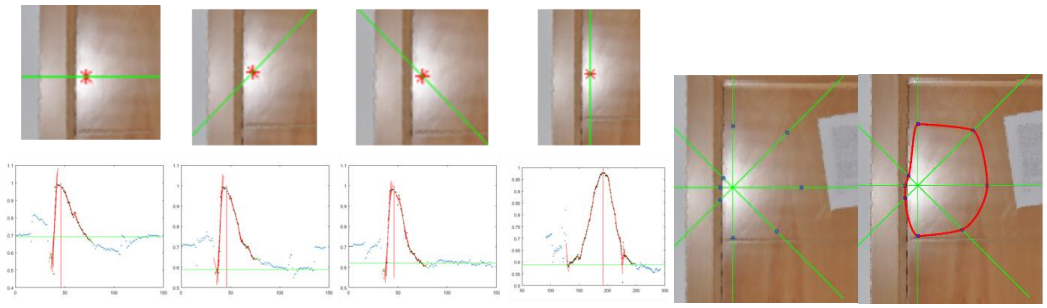


Figure 3. Detection of specular highlight regions in the wall. (a) Specular highlight Region of Interest (RoI). The RoI is marked in black in the figure on the right. (b) Specular Highlight Region Candidates are regions located within the RoI with peaks in the intensity image. (c) Specular region recognition by matching four intensity profiles to a Gaussian function (here the N/S profile fits a Gaussian function). (d) The extent of the specular region is defined by a spline whose control points are the eight end-points found in the intensity profiles where the slope reaches zero or the profile shows a significant discontinuity.

4.2 Colour merging

It is very frequent in our context for wall surfaces to be scanned from different viewpoints. But, the change in viewpoint leads to colour acquisitions with slightly different light conditions and responses, which can ultimately result in colour artefacts when naively merging the colour data. Specular highlights also negatively impact colouring during view merging. In order to provide a more realistic and homogeneous colouring to the wall, a weighted mean colour merging is proposed that considers the expected value (i.e. quality/reliability) of the colour data from each scanning location and the presence of specular highlights.

The first step of this process consists in correcting the specular highlights detected in each 4D orthoimage J_{CD}^{k} with the process described in Section 4.1. For this, we discard all colour information of the pixels contained in the highlight region and refill the region using the inpainting technique of Roth et al. [25]. Figure 4 shows an example of inpainting result obtained for three highlights, including the exemplar highlight of Figure 3.

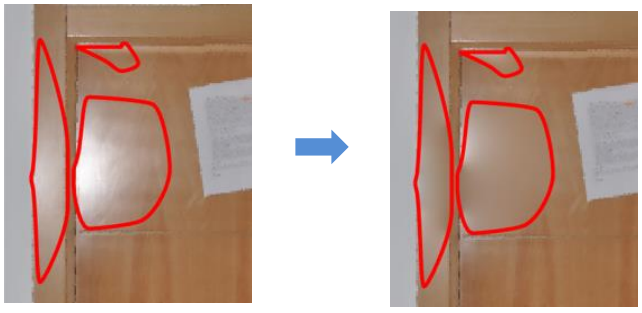


Figure 4 Repairing specular regions using inpainting. Result obtained for the exemplar highlight region of Figure 3.

Once specular highlight inpainting is complete, the colour information from the different views $\{J_{CD}^{k}\}$ is merged as follows. For each pixel P in the wall orthoimage, let $\{\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots \boldsymbol{c}_k\}$ be the k different colours obtained from the different scanning locations O_i, and let

$\{\theta_1, \theta_2, ..., \theta_k\}$ be the corresponding set of incidence angles. $\theta_i = (\vec{n}, \vec{u}_i)$, where $\vec{n}$ is the normal vector of the wall and $\vec{u}_i$ is the unitary vector of $\overrightarrow{PO_i}$. The merged colour assigned to *P*, *C(P)*, is formally calculated using the formula in Equation (1).

$$C(P) = \frac{\sum_{i=1}^{k} c_i w(\theta_i)}{\sum_{i=1}^{k} w(\theta_i)} \quad w \in \{w_1, w_2\} \tag{1}$$

where the weight *w* is picked from one of the two Gaussian functions w_1 or w_2, depending on the location of *P* in the wall.

w_1 is a Gaussian with mean μ_1=0 and standard deviation $\sigma_1=\pi/9$, whereas w_2 is a Gaussian with mean $\mu_2=\pi/9$ and standard deviation $\sigma_2=\pi/6$. Both are shown in Figure 5. If *P* does not lie in a specular region for the given position *k*, w_1 is chosen. This signifies: the more frontal the view, the higher the weight. The low σ_1 value is for making this criterion more exclusive. On the contrary, if *P* lies in a specular region, w_2 is used. This means that the colour of intermediate lateral views is considered more reliable than either frontal or very oblique ones. Note that, although the specular regions have been repaired, this merging strategy takes into account the still likely possibility of having a non-perfect filling result. A smoothing mean filter is finally applied over the contour of the highlights regions. Figure 6 shows the results obtained after merging three views of a wall.

The merging process leads to the creation of an orthoimage J_{CD} (with colour and depth information) of the wall where each pixel has the RGB components and the orthogonal distance to the wall plane.

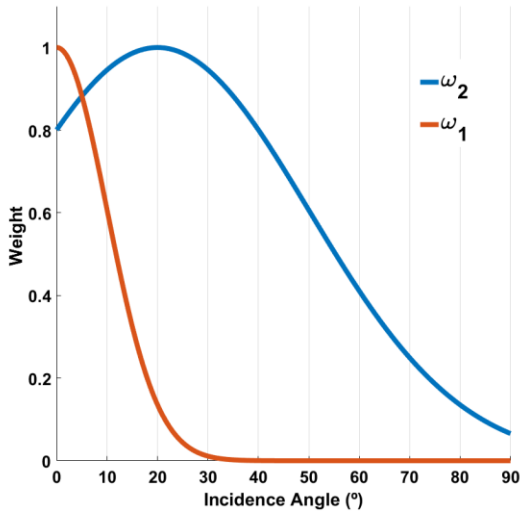


Figure 5. ω_1 and ω_2 gaussians.

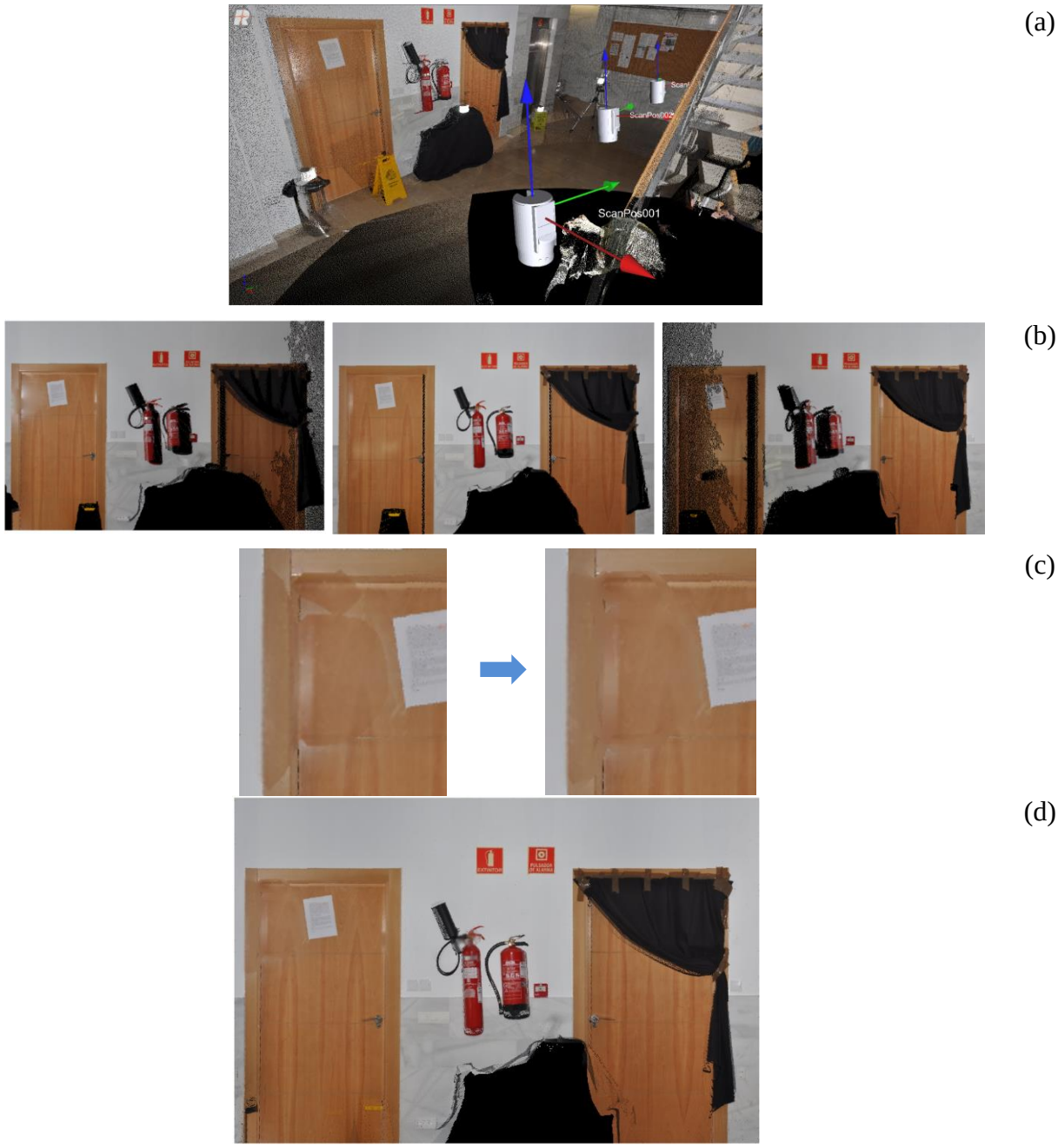


Figure 6. Illustration of the colour merging process. (a) Three different positions of the scanner capture data for one wall. (b) Set of wall orthoimages $\{J_{CD}^{k}\}$ generated from each of the three views. (c) Result of the colour merging on the highlight region (left) and border smoothing (right). (d) Result of the colour merging. Orthoimage J_{DC}.

5 DETECTION OF DOOR OPENINGS IN WALLS

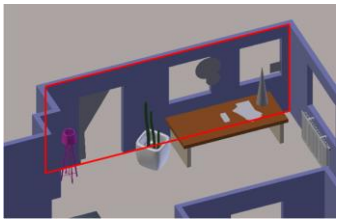
Our algorithm for recognising doors is defined under the following assumptions, which are true in the vast majority of cases:

- The walls are planar surfaces (this assumption is actually made at the SE recognition and modelling stage);
- Each wall has a fairly homogeneous colour but some variations may still exist, even after following the colour merging stage described above;
- Doors are rectangular with vertical and horizontal sides.

Door recognition is carried out in two stages, using a wall's labelled voxels and its orthoimage J_{CD}. The system first looks for *door openings* (i.e. openings that correspond to open or semi-open doors) as rectangular regions that contain *opening* voxels. This information is then used by the *door* recognition algorithm (see Section 6). This section explains how door openings are detected.

Opening voxels help to roughly localise various openings through the wall, that correspond to windows and doors (open or semi-open). In order to detect these, and particularly door openings, we employ the following five-step approach (see also illustration in Figure 7):

1. *Creation of trinary orthoimage I* (Figure 7 (b)). The trinary orthoimage *I* is generated. This has the same size as J_{CD} and its pixels are labelled *data* if the pixel is contained in an Occupied voxel, *centroid* if the pixel is contained in an Opening voxel and contains the centre of that voxel, and *no-data* otherwise (i.e. it is contained in an Opening or Occluded voxel).
2. *Extraction of a set of candidate rectangles in I* (Figure 7 (c)). Horizontal and vertical lines are found in *I* by using a lateral histogram algorithm [26]. All possible candidate rectangles defined by the intersections between pairs of vertical and horizontal lines are then computed.
3. *Centroid clustering* (Figure 7 (d)). The *centroids* in *I* are initially a set of dispersed pixels. Since these pixels represent opening regions, we employ a region growing algorithm to cluster these *centroids*.
4. *Best candidate rectangles* (Figure 7 (e)). For each cluster of *centroids*, the best candidate rectangle is selected as the rectangle with the smallest area that contains the largest number of *centroids*. We look for rectangles that contain *centroids* because the opening could be occluded, as in the example shown in Figure 7.
5. *Detection of door openings* (Figure 7 (e)). The rectangles with their lower side at the level of the floor are recognised as door openings (of open or semi-open doors).



(a)

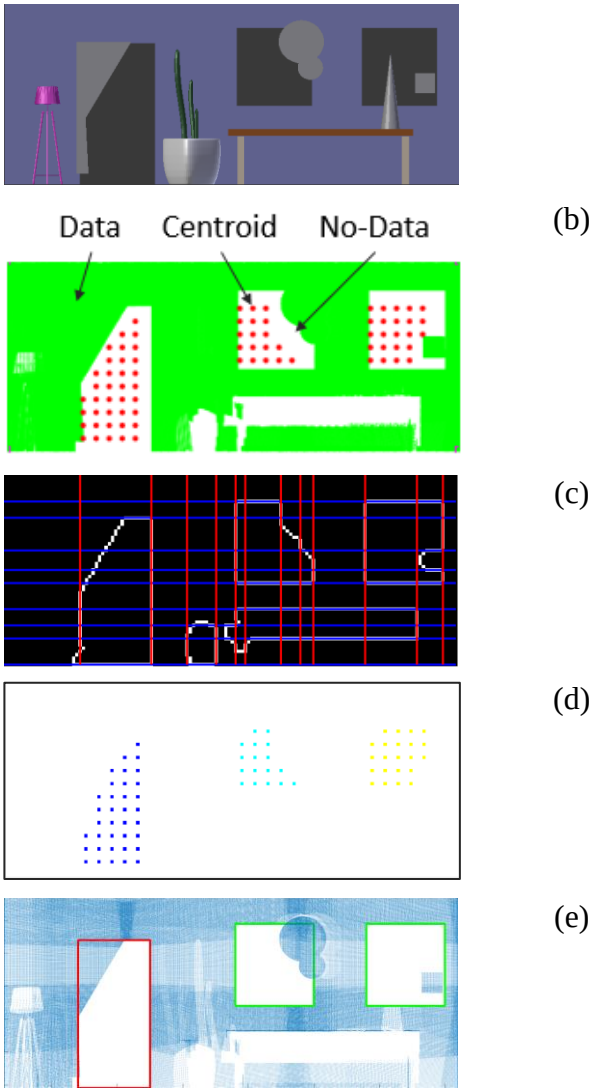


Figure 7. Detection of openings inside doors. (a) An example of wall with occluded openings. (b) The trinary image I composed of centroid (originating from the voxel space), data (3D point) and no-data (the lack of data). (c) Sets of horizontal (in blue) and vertical (in red) lines found in image I. (d) Groups of centroids after applying the region growing algorithm. (e) Three openings are detected with one, in red, detected as corresponding to a door (the other two would normally be detected as corresponding to windows).

6 DETECTION OF DOORS

The door detection algorithm detects and delimits the boundaries of the door for any state of the door (i.e. open, semi-open and closed). The output of the previous section yields essential information with which to classify the door. If the door contains an opening, the door is open or semi-open (depending on the opening angle), otherwise the door is closed. The opening angle is calculated using the set of points next to the door, as is explained in sub-section 6.3. We typically classify a rotating door leaf as an open door if its opening angle is equal to or greater than 90º; if the angle is below 90º, we classify the doors as semi-open. Closed doors do not contain openings and are typically co-planar with the wall plane. The method can also handle other kinds of doors, such as sliding doors, for which the opening

angle is 0º. In this case, the state of the door (open or semi-open) is inferred after comparing the opening size with the standard size of the sliding doors of the building.

To recognise doors, we have developed a 4D (colour + depth) approach that is able to deal with cases in which either the wall or the door do not have entirely uniform colours. Note that, although the proposed colour merging approach (Section 4) improves the uniformity of the colour information associated with the overall wall data, slight variations in colour may remain. As a result of this, the initial tests used to detect wall and door areas with simple colour thresholding algorithms yielded poor results, hence the proposed approach.

The algorithm for detecting doors is divided into two steps, wall area detection and door detection, described in the corresponding two sub-sections below.

6.1 Wall Area Detection

Taking J_{CD} as input, the segmentation of the visible parts of the wall is achieved as follows (with illustration in Figure 8, in which Figure 8 (a) shows the J_{CD} image).

1. *Finding coherent colour seeds* (Figure 8 (b)): First, small square patches (5×5 pixels), that we call 'seeds', are sampled regularly in J_{CD}. For each patch *m*, the distribution of the RGB-D pixel values {**v**} is analysed, and the patches for which the standard deviation σ in any of the four components is higher than a threshold (we use σ_{max} =0.2) are discarded. The process ensures that only patches that are *coherent* as regards both the colour domain and the depth (e.g. the patch is not located on the edge of frame) are retained.
2. *Clustering of coherent colour seeds* (Figure 8 (c) and Figure 8 (d)): Each *coherent* square patch *m* is then represented by the mean value of the RGB-D values of its 25 pixels $\bar{v}_m$ .An adaptive k-means algorithm is then employed to group the sample patches $\{\bar{v}_m\}$ into *k* clusters, where *k* is calculated by the algorithm itself Figure 8 (c)). The consistency within each cluster is then enhanced by removing any sample patch that has a *silhouette* value δ higher than a reasonable threshold ($|\delta|$>0.7) (Figure 8 (d)). The *silhouette* value for a member of a cluster is a measure, with a value of between -1 and 1, of how similar that member is to all the other members in the cluster, in comparison to members in the other.
3. *Wall area segmentation* (Figure 8 (e)): Finally, we find the set of pixels of J_{CD} associated with the the *i*-th cluster $\{v_m\}_{i\in[1;k]}$ by means of an exclusive thresholding matching technique imposed on all the four RGB-D components, and the wall area is recognised as the cluster that contains the largest number of pixels located on the left, right and top borders of the image.

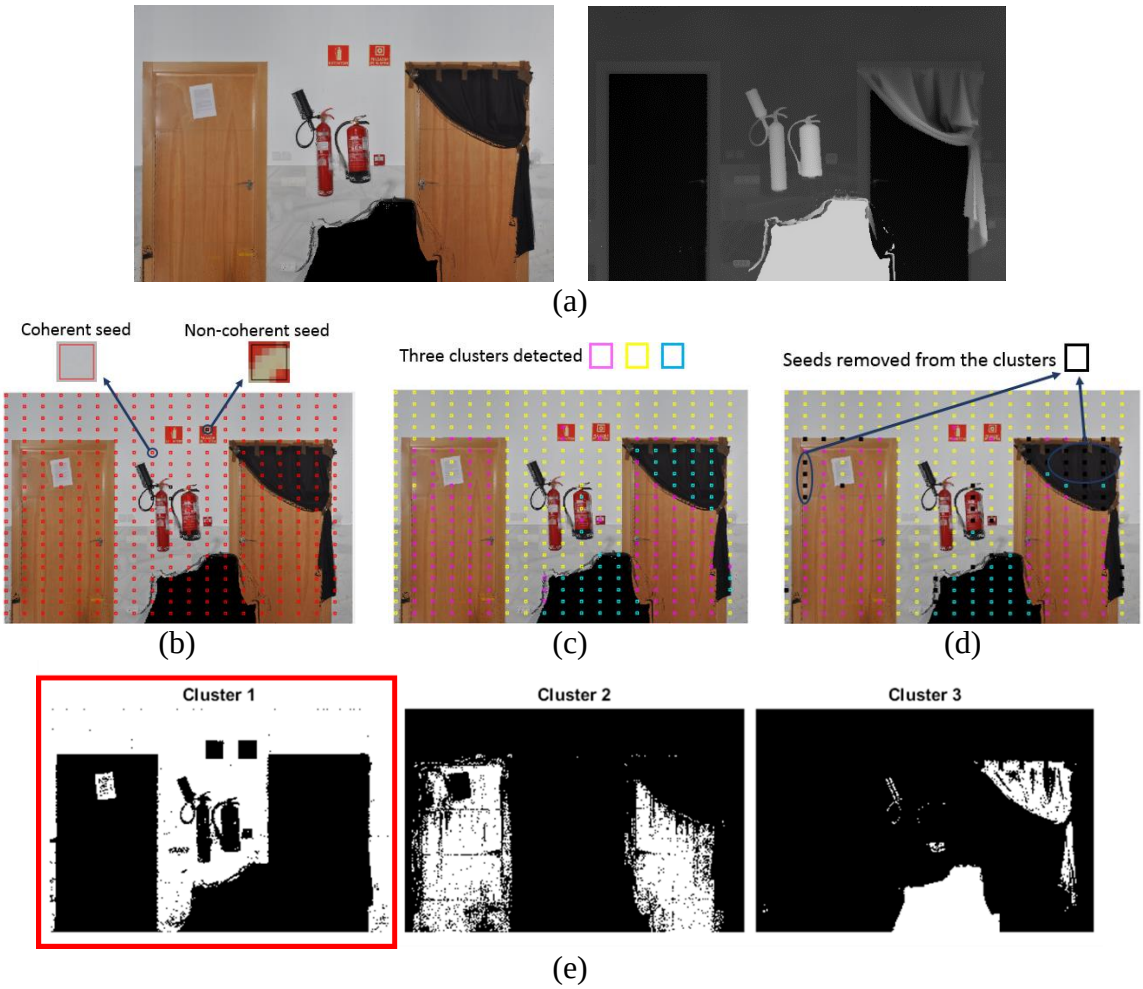


Figure 8. Wall area detection. (a) J_{CD} image with the colour component image J_C in the left image and the depth colour component J_D in the right image; (b) Detection of coherent 'seed' square patches; (c) Seed clusters. (d) Removal of inconsistent seeds from clusters. (e) The three clusters extracted in J_{CD} with the recognised wall area marked in red.

6.2 Door Detection

To recognise doors, we present an approach based on discontinuities in the 4D RGB-D space and the knowledge of the wall area. We process the colour and depth components of J_{CD} image separately, with J_{CD} decomposed into J_C (colour) and J_D (depth), and the results are finally recombined (see Figure 9). For J_C, a gradient operator is first applied to it that calculates the maximum change rate in the pixel colour (gradient) in the spectral dimensions [27]. This is followed by an image binarisation process, using Otsu's global histogram threshold technique that selects the threshold to minimize the intra-class variance of the black and white pixels. The result of this process is a binary image J'_C. For J_D (depth), the Canny edge detector is applied, generating a second binary image J'_D. J'_C and J'_D are finally combined using the OR operator to form a unified gradient image J'_{CD}.

White pixels in J'_{CD} represent discontinuities in the colour-depth space, which enables the detection of door frames as discontinuities in the colour domain only, in the depth dimension only, or in both. Given our assumption of rectangular door frames, we detect straight lines in J'_{CD} using the same approach as in Section 5 (Figure 10 (a)). These lines contain the colour-depth discontinuities of the wall (if the door has a protruding doorframe, the discontinuity in

the D dimension should result in line detections; if the door has a colour different from that of the wall, the discontinuity in the colour dimensions should also result in line detections). The detected lines contain *parts* of the contours of hypothetical doors. The word '*part*' is used here because occlusions may exist.

Next, similarly to the detection of door openings, we calculate all possible rectangles defined by two pairs of horizontal and vertical lines. Since we are looking for rectangles that delimitate doors, we only retain rectangles whose size falls within the range of typical door sizes and whose lowest edge lies at the bottom in image. This yields a highly reduced set of rectangles *{r}* (Figure 10 (b)). Each rectangle *r* is then recognised as an actual door if it fulfils the following conditions:

1. *Colour and depth consistency*: Using an adaptive k-means clustering process over the colour and depth data contained in rectangle *r*, the dominant colour and the dominant depth must both cover a certain percentage α_1 of the door area. We use α_1=50%.
2. *Door frame occlusion*: Each side of *r* must be supported by discontinuity information, i.e. white pixels in J'_{CD}, over at least the α_2 percentage of its length. We use α_2=60%, which means that 40% of occlusion of each side of the doorframe is permitted.
3. *Location consistency*: Not more than α_3=3% of the area enclosed by *r* intersects the wall area identified in the process described in Section 6.1.
4. *Minimum size*. *r* is not contained within any other rectangle that verifies conditions 1, 2 and 3. Note that in case of door openings, the minimum rectangle *r* always contains the opening's area.

Parameters α_1, α_2 and α_3 have been defined empirically. Section 8 is devoted to showing the influence of these parameters in the final result. Figure 10 (c) illustrates the final door recognition results. In this case, both doors are (correctly) recognized.

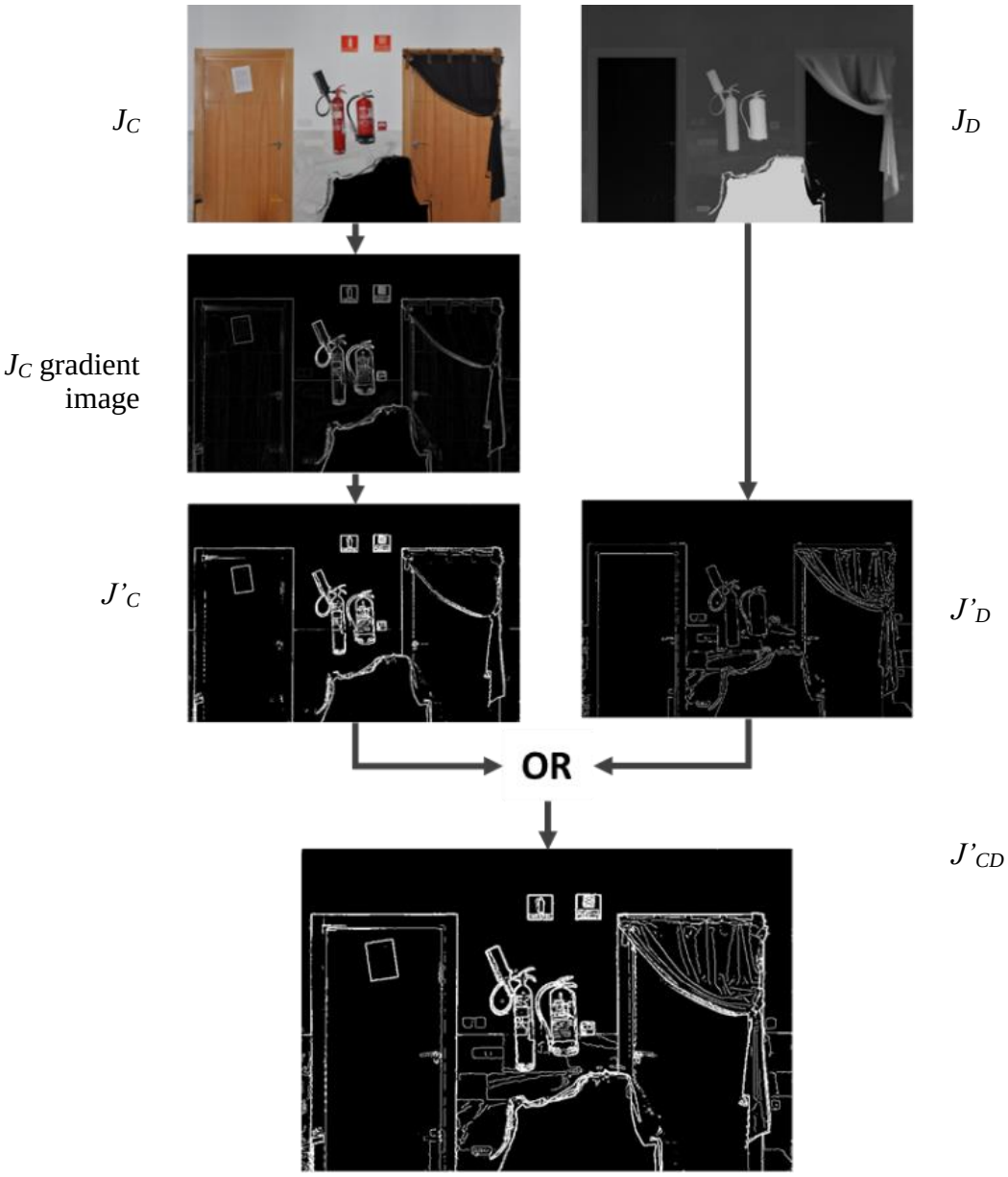
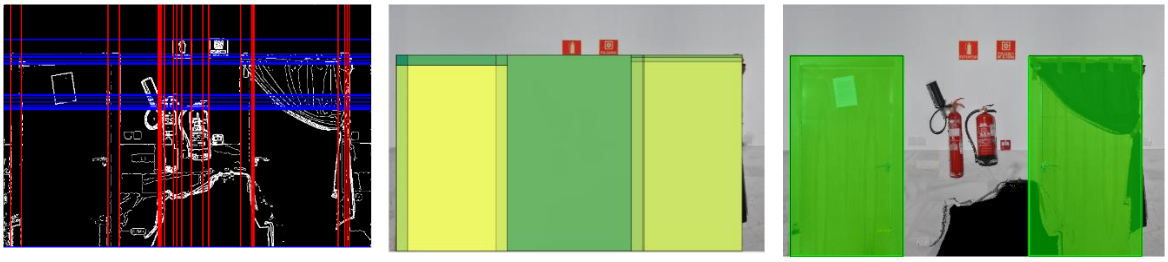


Figure 9. Generating the combined discontinuity image



(a) (b) (c)

Figure 10. Door detection example. In this case the algorithm recognizes two closed doors. (a) Horizontal and vertical lines are detected in the unified discontinuity image J'_{CD}. (b) Set of candidate rectangles {r}. (c) Recognized closed doors, with one presenting a significant level of occlusion.

6.3 Door Opening Angle

If the door contains an opening, the door's opening angle is obtained by taking a horizontal half-height splice of the door data and finding the line that best fits the points of the door leaf using RANSAC. Note that the line that represents the plane of the door is calculated with the door coordinates obtained in the previous step. In the case of rotating leaf doors, the angles are usually in the range [0º, 110º], whereas in that of sliding doors, the angles are zero or near zero.

Figure 11 illustrates an example of opening angles calculated for normal and sliding doors.

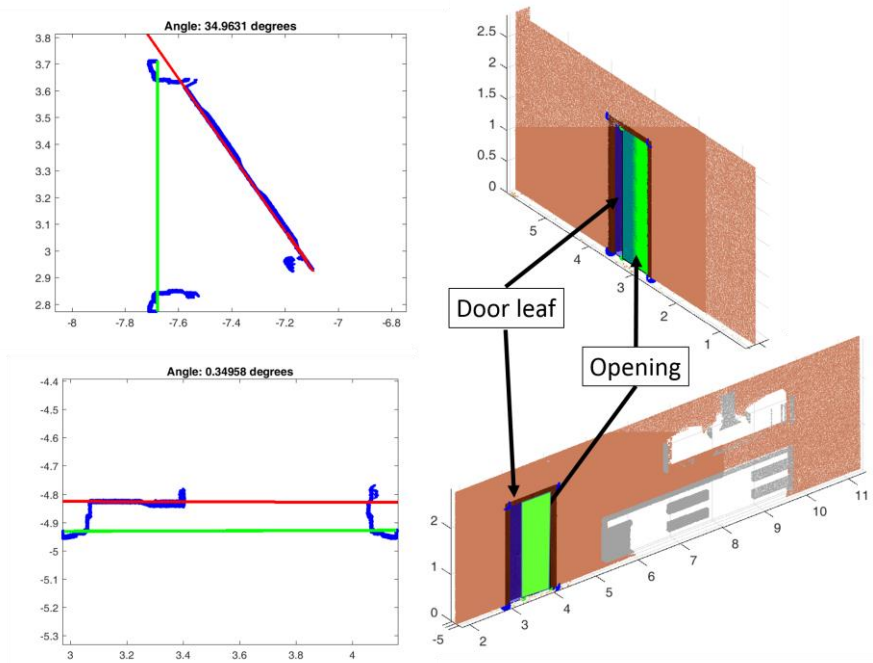


Figure 11 Calculation of opening angles. On the left, the red and green lines represent the wall plane and the door leaf. Top) A rotating leaf door. (Bottom) A sliding door. The opening is painted in transparent green and the door leaf is in blue.

7 EXPERIMENTAL VALIDATION

7.1 Performance Assessment Metrics

We evaluate the pose and size of the recognized doors by means of *Precision*, *Recall* and *F-measure* that are metrics frequently used in pattern recognition performance assessment. We compute these metrics based on the overlap between the areas of the ground truth (that is the correct door placed in the true position) and recognized doors. We evaluate the true-positive, false-positive and false-negative cases as follows (see Figure 12). Let Q and G be the areas of a pair of query and ground-truth doors. We define as true positive (t_p) the area of the detected door that is really a door, and false positive (f_p) the area of the detected door that does not belong to the ground-truth door. Finally, the false-negative (f_n) is defined as the area that belongs to a door but is not detected by our algorithm. Equations 2, 3 and 4 give the formal expression of t_p, f_p and f_n.

$$t_p = Q \cap G \tag{2}$$

$$f_p = Q - t_p \tag{3}$$

$$f_n = G - t_p \tag{4}$$

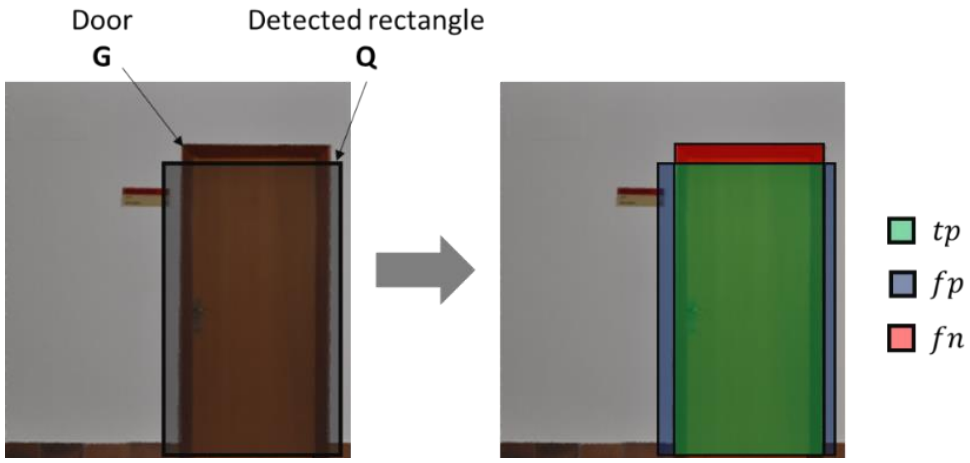


Figure 12. Definition of parameters t_p, f_p and f_n in openings.

Precision is defined as the fraction of the detected door's surface that is really a door (Equation 5), and *Recall* is the fraction of the door that is correctly recognized (Equation 6). *F-measure* (F_β) is a measure that combines *Precision* and *Recall*, using a kind of weighted average using a variable parameter β that defines whether more emphasis is put on *Precision* (i.e. the false detected door's area) or *Recall* (i.e. the undetected door's area) (Equation 7). Since, there is no clear argument in our context to prioritise precision over recall or vice versa, we report results for $\beta=0.5$, $\beta=1.0$ and $\beta=2.0$.

In order to give a more complete assessment of the performance of our method, we further introduce two measures for evaluating the error of the door model: the absolute global error (e_{abs}) and relative global error (e_{rel}) of a detected door (Equations 8 and 9). Note that these expressions, explained here for a single door, can be extended to all detected areas of the scene.

$$Precision = \frac{t_p}{f_p + t_p} \tag{5}$$

$$Recall = \frac{t_p}{f_n + t_p} \tag{6}$$

$$F_\beta = (1+\beta^2)\frac{Precision.Recall}{\beta^2.Precision + Recall} \tag{7}$$

$$e_{abs} = f_p + f_n \tag{8}$$

$$e_{rel} = \frac{f_p + f_n}{f_n + t_p} \tag{9}$$

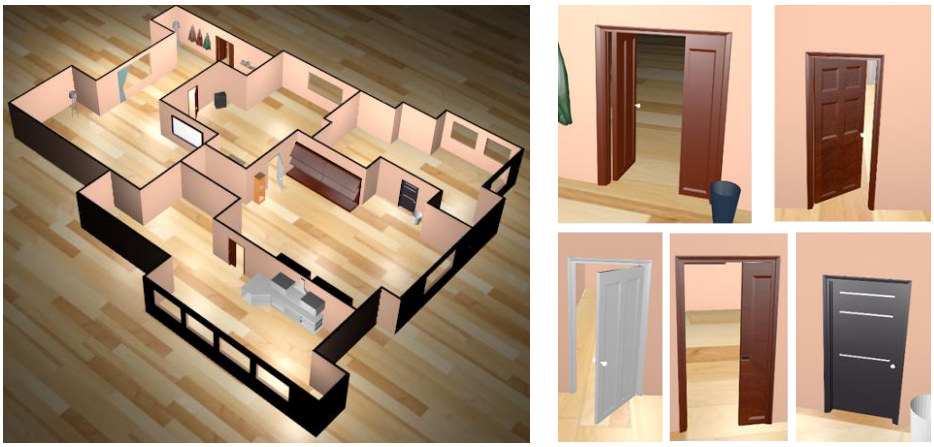
7.2 Experimental Dataset

In this paper, we present a new experimental dataset for door detection and modelling. The dataset is composed of coloured point clouds acquired in simulated and real environments.

7.2.1 Simulated Environment

The simulated scenario, illustrated in Figure 13, is the scanning of a 27m × 21m synthetic scene composed of 5 non-rectangular inhabited rooms, with 66 wall faces containing 5 doors with different opening angles. The synthetic model has been created with the Blender software, whereas the simulated scans have been obtained by using its add-on, Blensor [28], and following the scanning next-best-view scanning procedure developed for our automatic robotic system [24]. Blensor allows the simulation of scanning with a 3D laser scanner similar to ours, the Riegl VZ-400.

The simulation of real colour images is also possible with Blensor, signifying that the simulated data contains 3D coloured data, and can thus be used to test our door detection approach in very good simulation conditions. The advantage of the simulated data, however, is that the location of each door (and any other object in the environment) is known exactly. In other words, the ground truth data can be generated perfectly and automatically.



a) b)

Figure 13. a) 3D synthetic model in which the method has been tested. b) Details of doors with different opening angles.

7.2.2 Real Environments

Real data has been acquired with our robotic platform in real environments. The experimental robotic platform, called MoPAD (Mobile Platform for Autonomous Digitization), is composed of a Riegl VZ-400 3D laser scanner and a Nikon D90 camera on board a mobile robot (Robotnik Guardian). The robot is further equipped with two Hokuyo URG-04LX-UG01 sensors for autonomous navigation in buildings. The data contained in our shared dataset has been acquired in three different buildings of Castilla La Mancha University. Figure 14 shows some photos of the interiors tested.

In order to acquire the data, the mobile robot was manually moved to each room and the doors were then closed. Afterwards, our scan planning algorithm with an NBS strategy was executed. Data were also acquired for a few partially-closed doors to test the performance of

our approach in such cases. When the scanning process was complete, the accumulated point cloud was processed and the SEs (floor, ceiling and walls) extracted as described in Section 3. Finally, the 4D orthoimages were generated for each wall as described in Section 4. Ground truth models were built by manually selecting the vertices of the rectangles delimiting the doors in the 4D orthoimages.

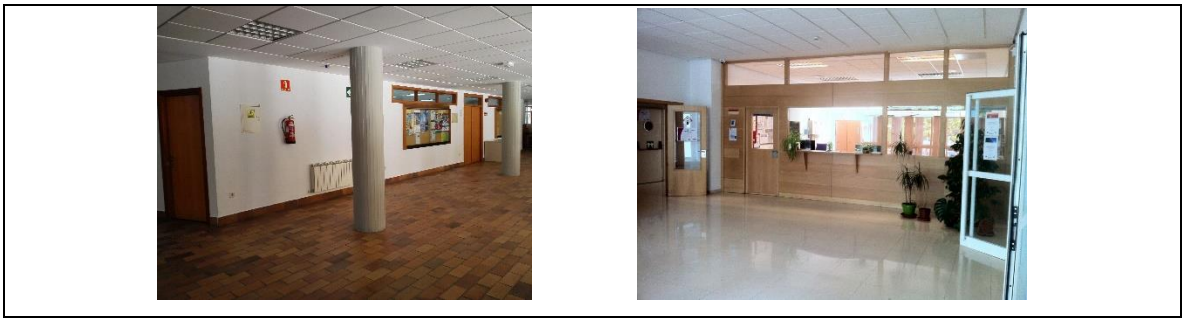


Figure 14. Views of the interiors of the three buildings considered to generate the real datasets.

The dataset is composed of coloured point clouds from 27 walls containing 35 doors. Figure 16 to Figure 20 illustrate several of the walls, showing various types of doors. Some of the walls also have windows and the majority contain other kinds of objects, either hung on the walls (e.g. papers, extinguishers, sockets, signs) or that partially occlude doors (e.g. curtains, posters). Different combinations of wall and door colours can be found, including very complex cases with doors co-planar to the wall and with a very similar colour to it, or walls that contain tiles with slightly different colours from the colour of the wall itself. Variations in the colour of the wall area were detected in the majority of cases. The range is between RGB variations of 5.49% (R), 5.49% (G) and 4.7% (B) in Figure 15 b), and variations of 27.84% (R), 29.41% (G), 27.84% (B) in Figure 20. Furthermore, the minimum RGB variation detected between the wall area and the door is 5.09% (R), 1.96% (G), 1.18% (B) in Figure 20.

The dataset has been generated under conditions of non-controlled illumination, except for the use of an automatic electronic camera flash, as explained earlier.

The wall scenes of which the shared dataset is composed can be classified into five categories:

1) *Simple scenes*. These are wall scenes (10 instances in the database) with no occlusion that contain one single or double door. Examples of such scenes are shown in Figure 16.

2) *Scenes with occlusions*. These are wall scenes (7 instances) that contain either of the two kinds of occlusions that can impact on the performance of our algorithm at two different stages: door panel occlusion and door frame occlusion. Examples of such scenes are shown in Figure 17.

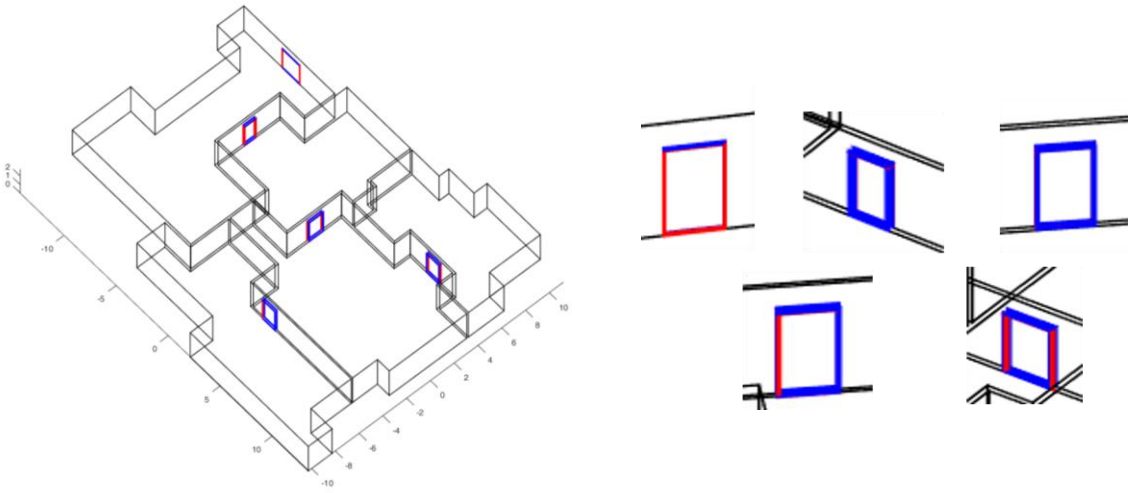
3) *Scenes with severe specular highlights*. Examples of such scenes are shown in Figure 18 (4 instances).

4) *Scenes with semi-open doors.* These are wall scenes (5 instances) in which one of the doors is neither fully closed nor fully open, but is partially open/closed to various levels. The shared dataset currently contains 3 doors with closing percentages from 1% to 70%. Examples of such scenes are shown in Figure 19.

5) *Very complex scenes*. These are wall scenes (9 instances) in which doors are co-planar to the wall or have a very similar colour to it. Examples of such scenes are shown in Figure 20, Figure 21 and Figure 23.

7.3 Door Detection in a Simulated Environment

Figure 15 shows the doors recognised in the simulated data (in red) superimposed onto the ground truth (in blue). The algorithm correctly detects all doors and there is only one false positive. In the case of the majority of the doors, a slight difference between the detected and ground truth rectangles is visible. This difference is evaluated using the statistics presented in Section 7.1,and all the results are summarised in Table 3. It can be seen that, in general, very high precision and recall values are achieved in the majority of cases (>0.95). The overall average values are 0.986 (Precision) and 0.983 (Recall). Furthermore, the average values of e_{abs} and e_{rel} are 0.077m^2 and 0.030 m^2 respectively. The latter signifies that, on average, the overlap error between the ground truth and detected door rectangles is only 8.1% of the surface of the doorframe rectangle. The *F-measure* values were $F_{0.5}$=0.9931, $F_{1.0}$=0.9905 and $F_{2.0}$=0.9880, which is a very good result (values above 0.85 are normally considered positive) and leads us to suggest that our approach could be used in cases in which either or both high recall and high precision are priorities. On the whole, these results suggest a good performance of our approach that achieves decent accuracy in the detection and localisation of doors.



a) b)

Figure 15. a) Door detection results. Rectangles of the ground truth are in blue and calculated rectangles are in red. b) Details of the doors detected.

Table 3. Evaluation of the results in the doors detection test.

Door	Op. angle	Precision	Recall	e_{abs} (m^2)	e_{rel}
#1	34,96°	0,993	0,989	0,035	0,017
#2	2,25°	0,990	0,991	0,065	0,018
#3	52,72°	0,953	0,953	0,239	0,092
#4	0°	0,987	0,983	0,058	0,028
#5	3,54°	0,995	0,986	0,045	0,017
#6	0°	0,987	0,987	0,066	0,025
#7	55,64°	0,990	0,982	0,069	0,026
#8	0°	0,987	0,986	0,067	0,026
#9	0,35°	0,994	0,986	0,049	0,018
Mean value	-	0,986	0,983	0,077	0,030

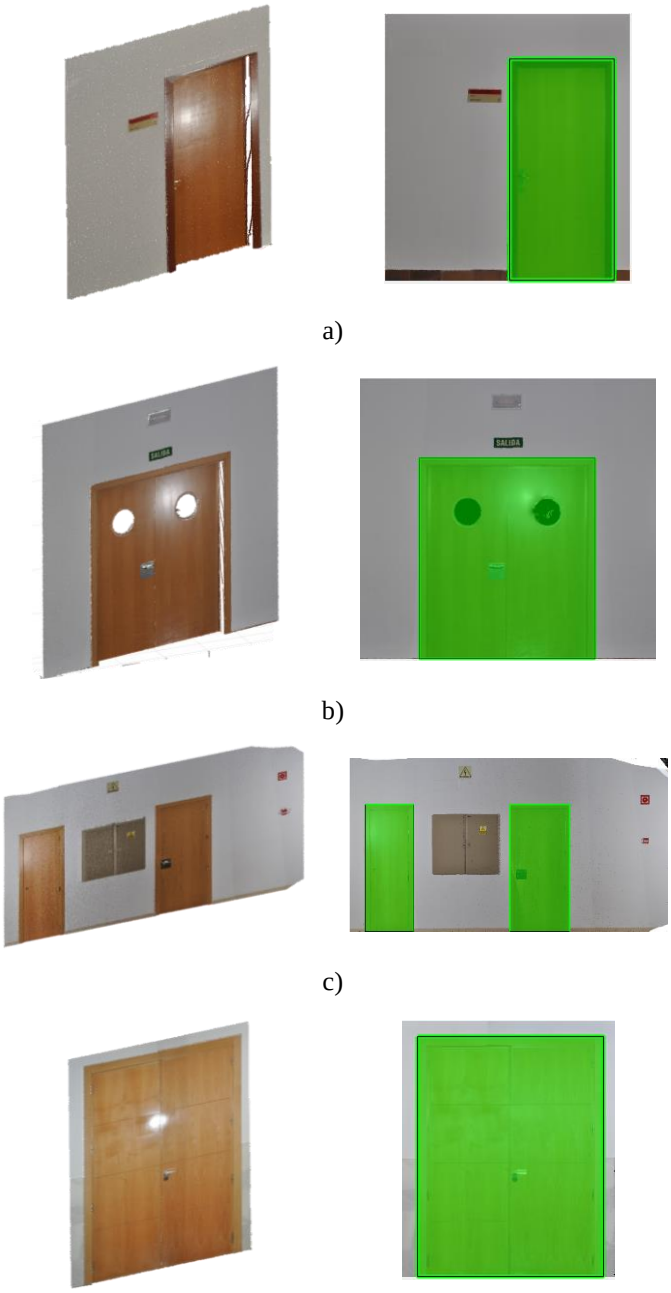
7.4 Door Detection in Real Environments

This section presents the results yielded by the algorithm in the case of closed doors and partially-closed doors, which are really the most interesting cases as regards accomplishing further robot interaction tasks, such as handle grasping and door opening.

The door detection algorithm successfully detects 34 of the 35 doors (97% detection rate) contained in the dataset acquired from real environments and yields two false positives in the case of complex walls (see Figure 20 and Figure 23). Figure 16 to Figure 20 present the results for a set of representative walls of the five wall classification categories:

1) *Simple scenes* (Figure 16). The method worked in all cases.
2) *Scenes with occlusions* (Figure 17). The results show that our method worked with door panel occlusion of up to 40%, and doorframe occlusion of up to 50%, which are unfortunately the highest occlusion levels in the current database.
3) *Scenes with severe specular highlights* (Figure 18). 56% of the walls led to significant specular highlights during scanning, owing to the smoothness of the surfaces involved. The algorithm correctly detected all those specular highlights, with small regions of about 8 cm^2, up to quite large regions of about 0.27 m^2.
4) *Scenes with semi-open doors* (Figure 19). The algorithm succeeded in all three cases.
5) *Very complex scenes* (Figure 20). Owing to the robustness of the procedure employed for calculating the wall area (explained in Section 6), the doors were successfully found in all but one of these cases. Figure 21 shows the intermediary results of the complete process in the particular complex scene in which the wall and doors are of more or less the same colour.

644

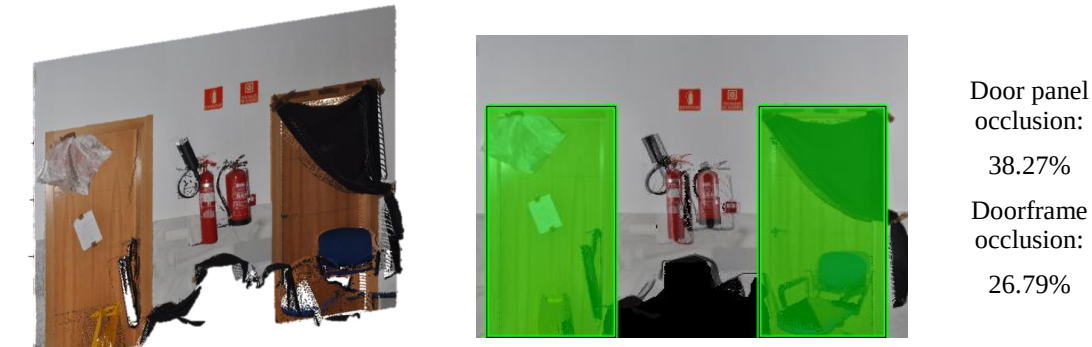


645 Figure 16. Detection results for Simple Scenes. (Left) Original 4D orthoimages. (Right) Door detection. Each
646 coloured rectangle represents a detected door.

647

a)

b)

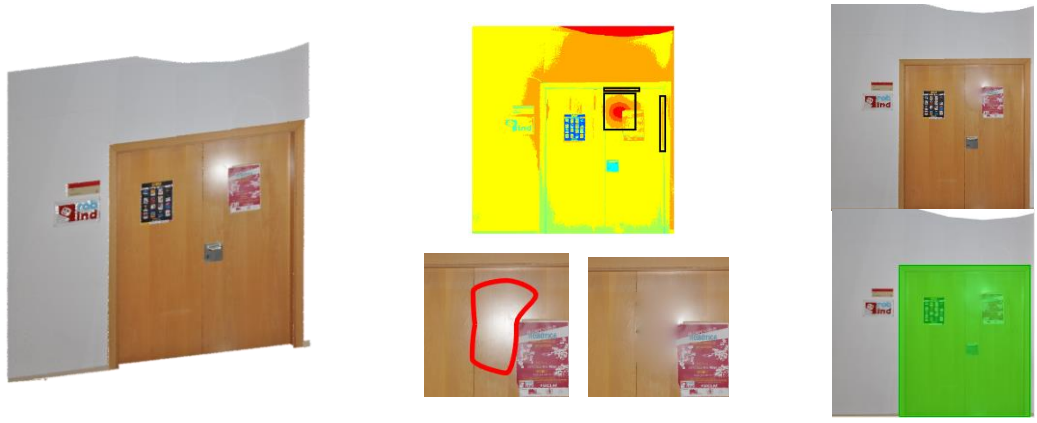


c)

d)

Figure 17. Detection results for Scenes with occlusion. Left: Original 4D orthoimages. Centre: Door detection. Right: Door and doorframe occlusion percentages.

a)



b)

c)

Figure 18. Detection results for scenes with severe specular highlights. Left: Original 4D orthoimages. Centre: Specular region detection and correction. Right: Corrected 4D orthoimages and door detection.

a)

b)

c)

d)

Figure 19. Detection results for Scenes with semi-open doors. Left: Original 4D orthoimages. Right: Door detection. Each coloured rectangle represents a detected door.

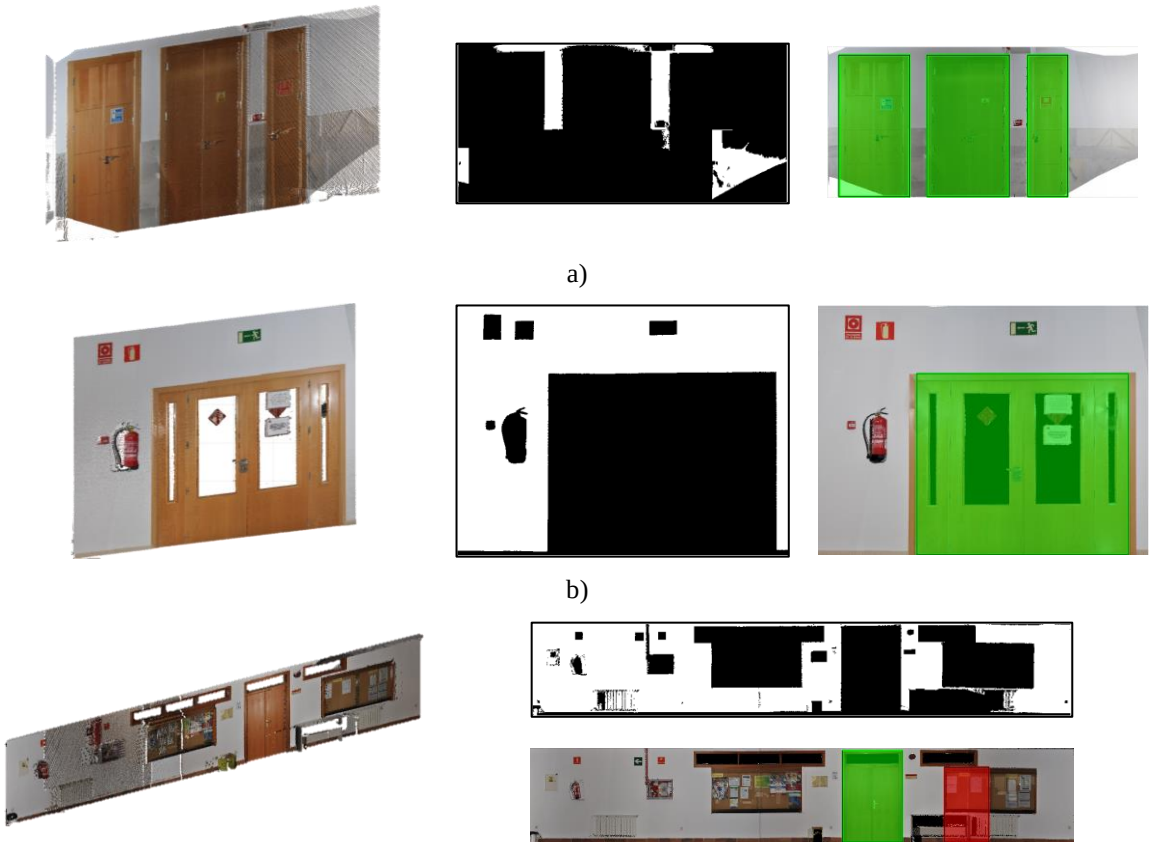


Figure 20. Detection results for Complex Scenes. Left: Original 4D orthoimages. Centre: Wall area detected in white. Right: Door detection. The last case shows a false positive.

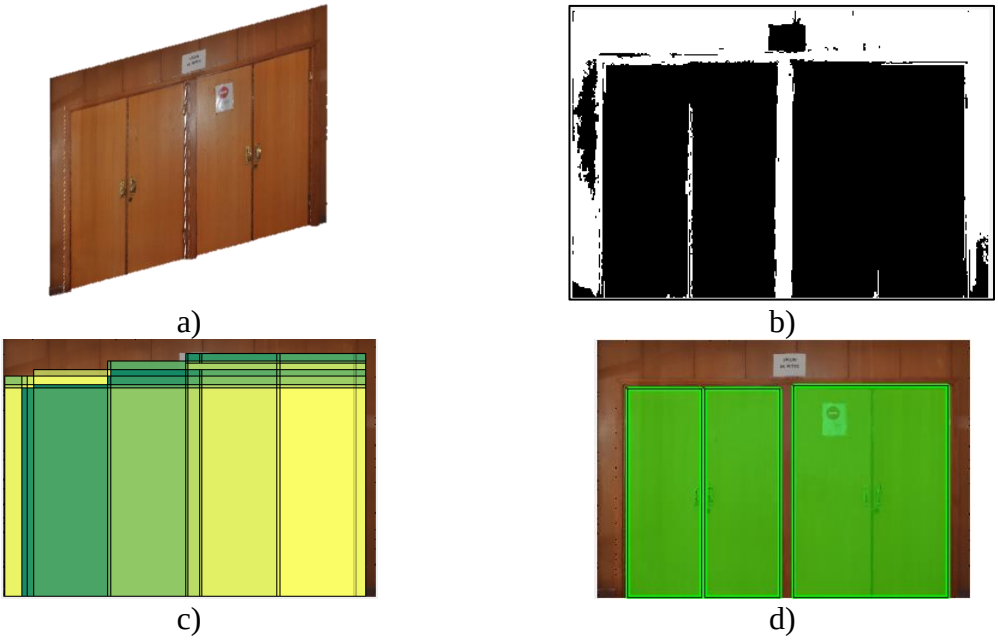


Figure 21. Intermediary results for a particularly complex case: a wooden double-door with a similar colour to the wall that is itself made up of multiple wooden panels. (a) View of the wall. (b) Wall area recognition. (c) Candidate rectangles superimposed in different colours. (d) Rectangles that enclose the detected doors.

The door positioning results are summarised in Table 4 that reports the mean precision, recall, absolute error, relative error and F_β calculated for ß =0.5, ß =1 and ß =2, for each of the five datasets sub-categories. Figure 22 additionally shows a precision-recall graph that includes all doors tested. In general, it can be stated that the method yields encouraging results. The average precision is above 98% in all cases, whereas recall values are in excess of 95%. For the error measures, the worst result is again for complex scenes with e_{rel}≈4,4 %, whereas e_{rel}≈4,1 % for all other scenes. F_β is above 0.95 for all cases and any choice of between ß =0.5, ß =1 and ß =2 does not provide any meaningful difference in the respective harmonic means. In other words, the approach performs equivalently whether recall or precision (or neither) is considered a priority. These results are encouraging, demonstrating the accuracy in the estimations of the size and position of the detected doors.

Table 4. Door positioning results. Mean values of Precision, Recall, Errors and F-measure for each of the five dataset sub-categories.

Cases	Instances	Precision	Recall	e_{abs}(m^2)	e_{rel}	$F_{0.5}$	F_1	F_2
Simple	10	0,995	0,993	0,032	0,012	0,997	0,992	0,988
Occlusion	7	0,989	0,971	0,115	0,041	0,999	0,997	0,996
Specular highlights	4	0,998	0,995	0,086	0,007	0,994	0,996	0,999
Semi-open doors	5	0,997	0,964	0,139	0,039	0,998	0,997	0,997
Complex walls	8	0,993	0,964	0,138	0,044	0,991	0,977	0,964

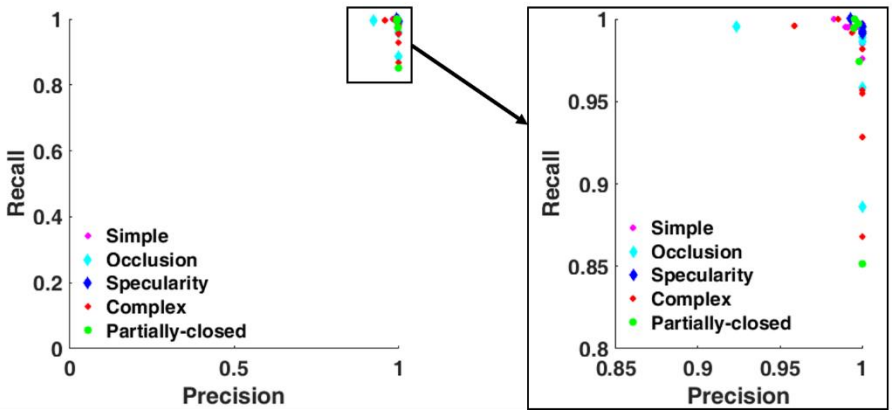


Figure 22. Precision vs. Recall graph

We shall now discuss cases in which our approach fails. The method fails mainly when the wall area and the doors are coplanar and are of a very similar colour. This occurs with the wall in Figure 23, in which the door is located on the left of the wall. The wall area detected is wrong (Figure 23 (d)) because the algorithm is not able to calculate the appropriate colour clusters. Thus, after removing inconsistent seeds from the clusters, the cluster that would normally best correspond to the wall (in yellow) remains incomplete, lacking the whole variety of colours of the wall and not including the edges of the orthoimages. As a result, the detection of the wall area is erroneous, and the validation of the rectangles does not yield the correct detection (false negative). In addition, owing to the occlusion on the bottom-right-hand part of the wall, the algorithm falsely detects a door (false positive case).

While this case provides a good illustration of the failure of our approach, it must be highlighted that these are particularly hard cases, and other existing detection methods would probably also fail.

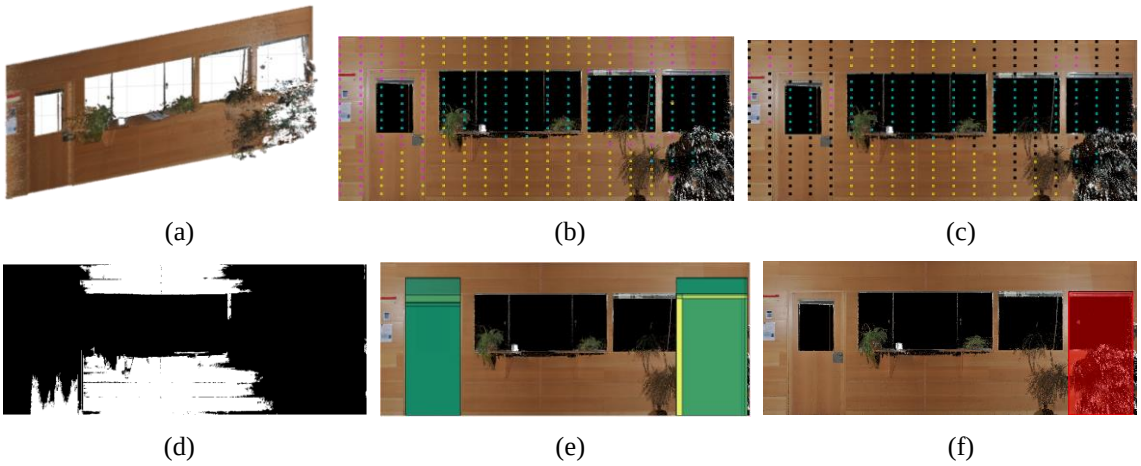


Figure 23. Example of situation in which our approach fails: a complex case in which the wall has a similar colour to that of the door, there are windows within door panels, and there are occlusions. The algorithm yields one false positive and one false negative. (a) 4D Orthoimage J_{CD}. (b) Set of initial square patches. (c) Seed clusters (rejected seeds in black). (d) Image of the detected wall area. (e) Set of candidate rectangles superimposed on the figure in various colours. (f) Final result.

7.5 Impact of the specular highlighting detection

Specular highlight detection is an important issue that increases the robustness of our proposal. Nevertheless, our approach could work with or without correction, depending on where specular highlighting takes place on the wall. In order to show the impact of specular highlights on the performance of our method, we have carried out an analysis of the method when this particular phase of the process is omitted.

Table 5 shows the results obtained for four significant walls on which the specular region can easily be seen. In this case, if the specular highlight detection is not carried out, the method fails. And in general, it always yields worse results (in terms of precision and recall). Basically, when the highlight falls between door and wall, the wall area is badly detected (see Section 6.1) and the method fails. This occurs with the wall in Figure 24. In this case, a part of the door is labelled as wall and the algorithm is not able to recognise the door. The conclusion is that, if the specular highlighting detection is not carried out, the method sometimes fails and always yields worst results (in terms of precision and recall).

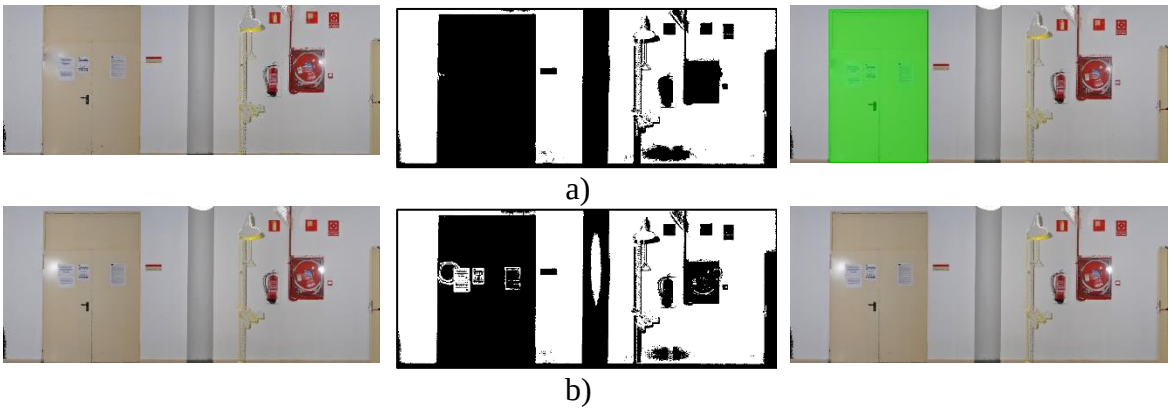


Figure 24 Results with (a) and without (b) specular highlight correction. From left to right: 4D orthoimage J_{CD}, wall area detected, door detection. The door is not detected without highlight correction.

Table 5 Results with and without specular highlight correction. (Walls: #1 (Figure 20 b)), #2 (Figure 18 b)), #3 (Figure 18 a)), #4 (Figure 18 c))

	Results with specular highlight correction					**Results without specular highlight correction**				
Walls	Doors detection	Precision	Recall	Absolute Error	Relative Error	Doors detection	Precision	Recall	Absolute Error	Relative Error
#1	Yes	1	0.928	0.390	0.071	Yes	1	0.918	0.443	0.081
#2	Yes	1	0.992	0.028	0.007	Yes	1	0.989	0.038	0.010
#3.1	Yes	1	0.991	0.019	0.008	Yes	0.9956	0.991	0.029	0.013
#3.2	Yes	1	0.995	0.010	0.004	Yes	1	0.995	0.010	0.004
#4	Yes	0.995	0.999	0.023	0.004	No	-	-	-	-

8 PARAMETER SELECTION

Our algorithm uses several parameters, which have been set by evaluating an independent subset of the data. We conducted a set of experiments with the aim of determining the effect of each parameter on the performance of the method. Table 6 shows the failure percentages and the minimum and mean precision and recall for each range of values tested. The main parameters are as follows.

- δ (Section 6.1). The silhouette value δ is necessary to refine colour seeds in the respective clusters. Table 6 shows that values up to 0.6 entail failures. On the other hand, excessive values of δ (i.e. 0.9) maintain the initial clustering and the inconsistent samples are not removed from the associated cluster. The threshold was, therefore, eventually set at 0.7.
- α_1 (Section 6.2). Parameter α_1 thresholds the dominant colour and depth of door candidates. Table 6 presents the performance of our method for a range of values of α_1. The main comment is that in the case of values greater than 0.5, the failures rate grows from 8% to 33%. Very low values, meanwhile, increase the risk of accepting candidate rectangles that contain a part of the wall or other objects next to the door.

As a result, we fixed the parameter at 50%. Figure 25 b) illustrates the results obtained for an example.

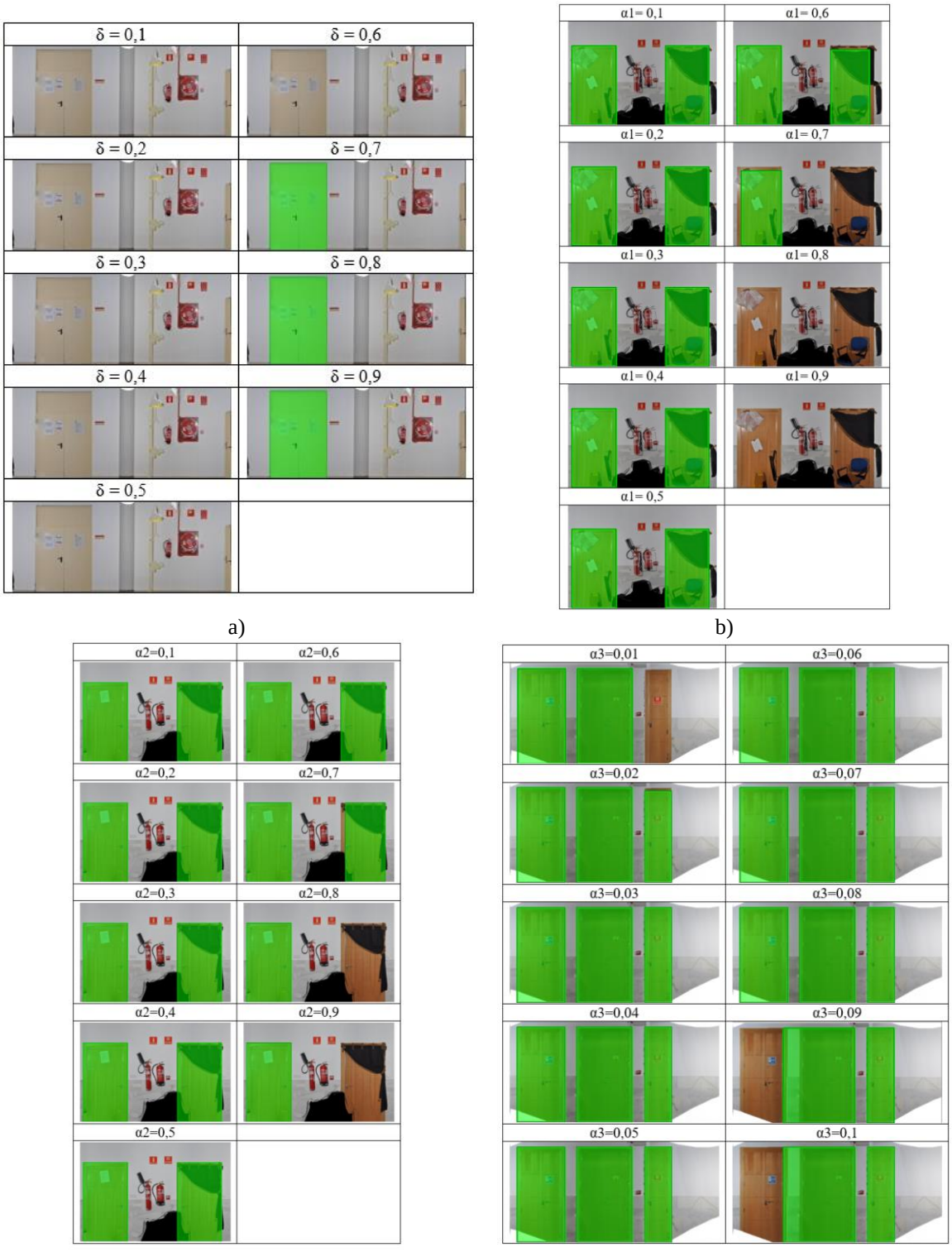
- α_2 (Section 6.2). Another interesting parameter related to occlusion is that of the percentage of door frame occlusion α_2. High values entail that the method will not admit a reasonable occlusion on the edges, as occurs for values greater than 0.7. On the contrary, in the case of low thresholds, many wrong candidate rectangles would be retained as candidates. According to the results obtained and reported for one example in Table 6, we have fixed this parameter at 60%. Figure 25 c) illustrates how the algorithm refuses best candidate rectangles with more occlusion for 0.7.

- α_3 (Section 6.2). In order to define the boundary of the door in a precise manner, parameter α_3 imposes the location consistency threshold. Table 6 shows that the thresholds that yield the best precision and recall values are 0.03 or 0.04. In order not to run the risk of introducing bigger errors in the door size and make the method more sensitive to local noise and small imprecisions in the wall area detected, we set this parameter at 0.03 (see Figure 25 d)).

- σ (Section 6.1). We use the standard deviation σ of the four components (RGB+depth) to find coherent colour seeds in the RGB image. In this case, 0.2 was found to be the most suitable value.

Table 6 Results of the door detection method for different values of δ, α_1, α_2 and α_3.

	Value	0,1 * p	0,2 * p	0,3 * p	0,4 * p	0,5 * p	0,6 * p	0,7 * p	0,8 * p	0,9 * p
δ (p=1)	Failures	20%	20%	20%	20%	20%	20%	**0%**	0%	20%
	Min. P.	0,958	0,958	0,958	0,958	0,958	0,958	**0,958**	0,958	0,779
	Min. R.	0,981	0,981	0,981	0,981	0,981	0,981	**0,981**	0,981	0,988
	Mean P.	0,99	0,99	0,99	0,99	0,99	0,99	**0,992**	0,992	0,937
	Mean R.	0,988	0,988	0,988	0,988	0,988	0,988	**0,992**	0,992	0,994
α_1 (p=1)	Failures	0%	0%	0%	0%	**0%**	8%	17%	25%	33%
	Min. P.	0,924	0,924	0,924	0,924	**0,924**	0,924	0,869	0,959	0,959
	Min. R.	0,872	0,872	0,872	0,872	**0,872**	0,496	0,496	0,462	0,433
	Mean P.	0,985	0,985	0,985	0,985	**0,985**	0,985	0,978	0,996	0,996
	Mean R.	0,976	0,976	0,976	0,976	**0,976**	0,899	0,751	0,728	0,754
α_2 (p=1)	Failures	10%	0%	0%	0%	0%	**0%**	0%	10%	10%
	Min. P.	0,485	0,858	0,87	0,901	0,901	**0,924**	0,924	0,959	0,983
	Min. R.	0,619	0,978	0,978	0,978	0,978	**0,978**	0,894	0,889	0,462
	Mean P.	0,931	0,97	0,972	0,977	0,977	**0,983**	0,983	0,995	0,997
	Mean R.	0,961	0,992	0,992	0,992	0,992	**0,991**	0,975	0,976	0,866
α_3 (p=0,1)	Failures	17%	0%	**0%**	**0%**	0%	0%	0%	0%	17%
	Min. P.	0,959	0,959	**0,959**	**0,959**	0,933	0,933	0,933	0,933	0,779
	Min. R.	0,966	0,963	**0,982**	**0,982**	0,982	0,982	0,982	0,982	0,988
	Mean P.	0,99	0,99	**0,99**	**0,99**	0,968	0,968	0,968	0,968	0,935
	Mean R.	0,985	0,988	**0,992**	**0,992**	0,994	0,994	0,994	0,994	0,994



761 Figure 25 Examples of doors detected for different values of parameter a) δ, b) α_1, c) α_2 and d) α_3.

9 CONCLUSIONS AND FUTURE WORK

This paper presents an integrated approach for the detection, localisation and sizing of doors that are either closed or open. The detection is carried out in coloured 3D laser scanned point clouds. The detection of open doors is based on the detection of rectangular data holes in the wall planes, while the detection of closed doors is based on the detection of the actual wall area and the subsequent processing of the rectangular areas not corresponding to the wall. This unique approach can handle occlusion and uses both 3D geometry and colour for more robust detections and localisations. Its robustness and performance are validated experimentally using a dataset of simulated and real data (including ground-truth information) from wall scenes of various complexities. A dataset [23] composed of 19 coloured point clouds corresponding to real walls that contain at least one door is made publically available to the research community. The experimental evaluation shows that the proposed approach works in very challenging cases in which doors are closed, are co-planar to the wall, or/and are of a very similar colour to it.

Future work will focus on addressing more complex cases (some of which are contained in the shared datasets). We shall particularly focus on developing a more robust algorithm for the detection of wall areas that have significant colour variations and on approaches for non-rectangular doors. With regard to the first issue, we are working on the segmentation and integration of different colour-coherent parts of wall areas. With regard to the second subject, we assume that non-rectangular doors are essentially those with rounded archways. We hope that a matching technique may be employed to recognise different archways in the image J'_{CD}.

ACKNOWLEDGMENTS

This work was supported by the Spanish Economy and Competitiveness Ministry (DPI2013-43344-R project, AEI/FEDER, UE) and by the Castilla La-Mancha Government (PEII-2014-017-P project).

REFERENCES

[1] D. Dai, G. Jiang, J. Xin, X. Gao, L. Cui, Y. Ou, G. Fu, Detecting, locating and crossing a door for a wide indoor surveillance robot, 2013 IEEE Int. Conf. Robot. Biomimetics, ROBIO 2013. (2013) 1740–1746. doi:10.1109/ROBIO.2013.6739719.

[2] A.H. Adiwahono, Y. Chua, K.P. Tee, B. Liu, Automated door opening scheme for non-holonomic mobile manipulator, Int. Conf. Control. Autom. Syst. (2013) 839–844. doi:10.1109/ICCAS.2013.6704030.

[3] P. Tang, D. Huber, B. Akinci, R. Lipman, A. Lytle, Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques, Autom. Constr. 19 (2010) 829–843. doi:10.1016/j.autcon.2010.06.007.

[4] X. Yang, Y. Tian, Robust door detection in unfamiliar environments by combining edge and corner features, 2010 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. - Work. CVPRW 2010. (2010) 57–64. doi:10.1109/CVPRW.2010.5543830.

[5] M.M. Shalaby, M.A.M. Salem, A. Khamis, F. Melgani, Geometric model for vision-based door detection, Proc. 2014 9th IEEE Int. Conf. Comput. Eng. Syst. ICCES 2014. (2015) 41–

46. doi:10.1109/ICCES.2014.7030925.

[6] A. Andreopoulos, J.K. Tsotsos, Active vision for door localization and door opening using playbot: A computer controlled wheelchair for people with mobility impairments, Proc. 5th Can. Conf. Comput. Robot Vision, CRV 2008. (2008) 3–10. doi:10.1109/CRV.2008.23.

[7] W. Chen, T. Qu, Y. Zhou, K. Weng, G. Wang, G. Fu, Door recognition and deep learning algorithm for visual based robot navigation, 2014 IEEE Int. Conf. Robot. Biomimetics, IEEE ROBIO 2014. (2014) 1793–1798. doi:10.1109/ROBIO.2014.7090595.

[8] S. Kim, H. Cheong, D.H. Kim, S.K. Park, Context-based object recognition for door detection, IEEE 15th Int. Conf. Adv. Robot. New Boundaries Robot. ICAR 2011. (2011) 155–160. doi:10.1109/ICAR.2011.6088578.

[9] R. Sekkal, F. Pasteau, M. Babel, B. Brun, I. Leplumey, Simple monocular door detection and tracking, 2013 IEEE Int. Conf. Image Process. ICIP 2013 - Proc. (2013) 3929–3933. doi:10.1109/ICIP.2013.6738809.

[10] L.C. Goron, L. Tamas, G. Lazea, Classification within indoor environments using 3D perception, Autom. Qual. Test. Robot. (AQTR), 2012 IEEE Int. Conf. (2012) 400–405. doi:10.1109/AQTR.2012.6237743.

[11] S.M.Z. Borgsen, M. Schöpfer, L. Ziegler, S. Wachsmuth, Automated door detection with a 3D-sensor, Proc. - Conf. Comput. Robot Vision, CRV 2014. (2014) 276–282. doi:10.1109/CRV.2014.44.

[12] K.M. Varadarajan, M. Vincze, 3D room modeling and doorway detection from indoor stereo imagery using feature guided piecewise depth diffusion, IEEE/RSJ 2010 Int. Conf. Intell. Robot. Syst. IROS 2010 - Conf. Proc. (2010) 2758–2765. doi:10.1109/IROS.2010.5651525.

[13] T.H. Yuan, F.H. Hashim, W.M.D.W. Zaki, A.B. Huddin, An Automated 3D Scanning Algorithm using Depth Cameras for Door Detection, 2015 Int. Electron. Symp. (2015) 58–61. doi:10.1109/ELECSYM.2015.7380814.

[14] M. Derry, B. Argall, Automated doorway detection for assistive shared-control wheelchairs, Proc. - IEEE Int. Conf. Robot. Autom. (2013) 1254–1259. doi:10.1109/ICRA.2013.6630732.

[15] L. Díaz-Vilariño, K. Khoshelham, J. Martínez-Sánchez, P. Arias, 3D Modeling of Building Indoor Spaces and Closed Doors from Imagery and Point Clouds, Sensors. 15 (2015) 3491–3512. doi:10.3390/s150203491.

[16] N. Banerjee, X. Long, R. Du, F. Polido, S. Feng, C.G. Atkeson, M. Gennert, T. Padir, Human-Supervised Control of the ATLAS Humanoid Robot for Traversing Doors, IEEE-RAS Int. Conf. Humanoid Robot. (2015) 722–729. doi:10.1109/HUMANOIDS.2015.7363442.

[17] X. Xiong, A. Adan, B. Akinci, D. Huber, Automatic creation of semantically rich 3D building models from laser scanner data, Autom. Constr. 31 (2013) 325–337. doi:10.1016/j.autcon.2012.10.006.

[18] J. Xu, K. Kim, L. Zhang, D. Khosla, 3D Perception for Autonomous Robot Exploration, Int. Symp. Vis. Comput. (2015) 888–900. doi:10.1007/978-3-319-27857-5 79.

[19] A. Budroni, J. Böhm, Automatic 3D Modelling of Indoor Manhattan-World Scenes From Laser Data, XXXVIII (2010).

[20] B. Kakillioglu, K. Ozcan, S. Velipasalar, Doorway detection for autonomous indoor navigation of unmanned vehicles, IEEE Int. Conf. Image Process. (2016) 3837–3841. doi:10.1109/ICIP.2016.7533078.

[21] R.B. Rusu, Z.C. Marton, N. Blodow, M. Dolha, M. Beetz, Towards 3D Point cloud based object maps for household environments, Rob. Auton. Syst. (2008). doi:10.1016/j.robot.2008.08.005.

[22] B. Quintana, S.A. Prieto, A. Adán, F. Bosché, Door Detection in 3D Colored Laser Scans for Autonomous Indoor Navigation, 2016 Int. Conf. Indoor Position. Indoor Navig. (2016).

[23] 3D Visual Computing & Robotics Lab, (n.d.). http://isa.esi.uclm.es/.

[24] B. Quintana, S.A. Prieto, A. Adán, A.S. Vázquez, Semantic Scan Planning for Indoor Structural Elements of Buildings, Adv. Eng. Informatics. (2016).

[25] S. Roth, M.J. Black, Fields of experts: A framework for learning image priors, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2 (2005) 860–867. doi:10.1109/CVPR.2005.160.

[26] E.R. Davies, Machine Vision: Theory, Algorithms, Practicalities., Morgan Kaufmann Publishers Inc. San Francisco, CA, USA, 2004.

[27] S. Di Zenzo, A note on the gradient of a multi-image, Comput. Vision, Graph. Image Process. 33 (1986) 116–125. doi:10.1016/0734-189X(86)90223-9.

[28] M. Gschwandtner, R. Kwitt, A. Uhl, W. Pree, BlenSor: Blender sensor simulation toolbox, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 6939 LNCS (2011) 199–208. doi:10.1007/978-3-642-24031-7_20.